11-3-2003

Rendering HDR images

Garrett Johnson

Mark Fairchild

Rendering HDR Images

Garrett M. Johnson and Mark D. Fairchild
Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology
Rochester, New York, USA

Abstract

Color imaging systems are continuously improving, and have now improved to the point of capturing high dynamic range scenes. Unfortunately most commercially available color display devices, such as CRTs and LCDs, are limited in their dynamic range. It is necessary to tone-map, or render, the high dynamic range images in order to display them onto a lower dynamic range device. This paper describes the use of an image appearance model, iCAM, to render high dynamic range images for display. Image appearance models have greater flexibility over dedicated tone-scaling algorithms as they are designed to predict how images perceptually appear, and not designed for the singular purpose of rendering. In this paper we discuss the use of an image appearance framework, and describe specific implementation details for using that framework to render high dynamic range images.

Introduction

In everyday life we encounter a huge range of absolute luminance levels, most of which the visual system handles with ease. Perhaps more impressive is the visual systems ability to instantaneously and seamlessly adapt to scenes with a large dynamic range, scenes that can exceed 10000 to 1 between sunlight and shadows.

Recent advances in color imaging have lead to systems that are capable of capturing these high dynamic range (HDR) scenes. These systems can be based upon multiple photographic exposures, as described by Debevec[1] and Xiao *et al,*[2] or sensor improvements that make it capable to capture high dynamic range information with a single exposure.[3,4,5] Likewise, these systems might be high-contrast computer graphics renderings as described by Ward *et al.*[6,7]

While the systems for capturing high dynamic range images have improved over the years, the systems for displaying these images have not kept up. A typical desktop display is capable of displaying only one or two orders of magnitude of dynamic range. As such, there has been much research into developing tone reproduction algorithms for rendering high dynamic range images onto lower dynamic range displays. Detailed reviews of many of these algorithms can be found in Reinhard[8], Durand[9], DiCarlo[10] and Pattanaik.[11,12] Many of these algorithms do an admirable job of rendering HDR scene for display on low-dynamic range display devices. In general these algorithms are designed for a single purpose: rendering high dynamic range scenes onto lower dynamic range displays. The techniques used vary, though most are based at some level on a perceptual model of human contrast (either local or global). These techniques can typically be thought of as perceptually based image processing models.

There have been more comprehensive models of the human visual system that have been used to guide tone reproduction of high dynamic range images. One example of this type of model is the multiscale observer model, described by Pattanaik *et al.*[11,12] The spatial and chromatic adaptation facilities of this model were capable of tone-scaling high dynamic range images, though the model was not limited to just this purpose. This paper describes a similar approach to tone mapping, through the use of an image appearance model. The image appearance model proves to be quite adept at rendering high dynamic range images, though that is just one design goal.

Image Appearance Model

A next generation image appearance model, coined iCAM, was recently introduced.[13] Some of the defined goals in formulating iCAM were to combine traditional color appearance capabilities along with spatial vision and image quality metrics. iCAM was designed to be computationally simpler than the multiscale observer model, with similar capabilities.

Image appearance models attempt to predict the perceptual response towards spatially complex stimuli. As such, they can provide a unique framework for the prediction of the appearance of high dynamic range images. It is important to stress that these models, such as iCAM, are not designed specifically as tone-mapping algorithms but rather as predictors of overall color appearance. However, the general iCAM framework does not need to be changed to be useful for the rendering of high dynamic range scenes. Several of the parameters, or modules, of the iCAM framework can be specifically tuned for this application, just as they can be tuned for image difference and quality predictions.[13] This paper focuses on one specific implementation of the iCAM framework for high dynamic range tone mapping, while examining the effects of several of the individual parameters.

Figure 1 shows the general flowchart of the iCAM image appearance model, including the inverse model used for displaying high dynamic range images.

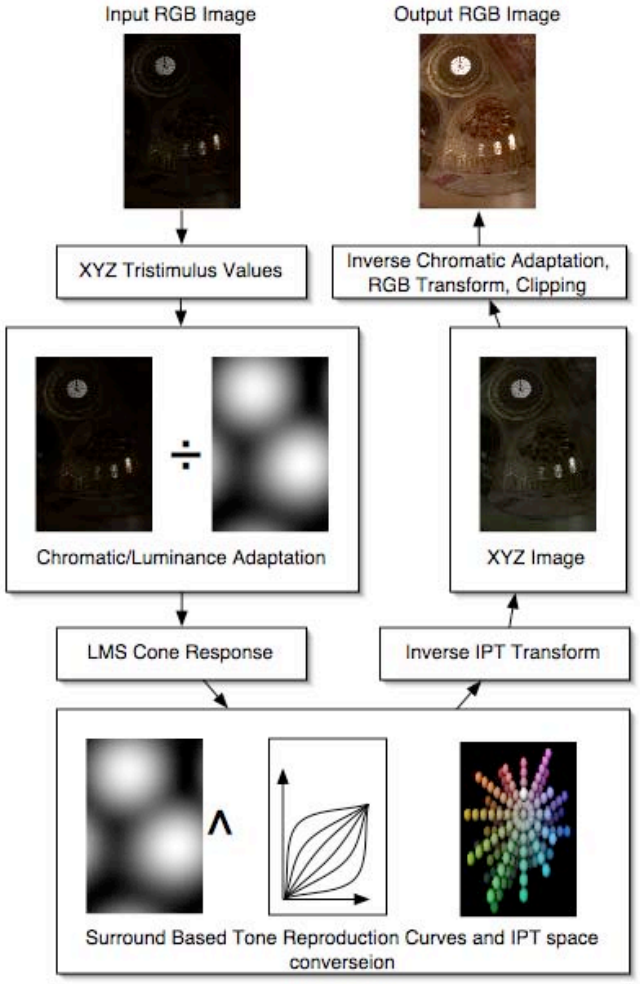

Figure 1. Flow chart of iCAM for rendering high dynamic range images.

The input into the model is a high dynamic range image, typically a floating point RGB image. Ideally the input would be a characterized linear RGB, CIE XYZ, or spectral image. The RGB image is then transformed into CIE 1931 XYZ tristimulus values, through device characterization (or assumption if necessary). An example transformation using the sRGB color space is described in Equation 1.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} .4124 & .3576 & .1805 \\ .2126 & .7152 & .0722 \\ .0193 & .1192 & .9505 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Chromatic Adaptation

Once the input image is in device independent coordinates, the next stage is the chromatic and luminance adaptation transform. This step serves two purposes: local adaptation of the high dynamic range scenes, and a global whitepoint shift towards CIE D65. This whitepoint shift is necessary, as the IPT color space used in iCAM is defined only for D65.[14] The chromatic adaptation transform is identical to that of CIECAM02, which is a linear von Kries transformation with an incomplete adaptation factor.[15] This transform is shown in Equations 2-4.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathbf{M}_{CAT02} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (2)$$

$$\mathbf{M}_{CAT02} = \begin{bmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.0136 & 0.9834 \end{bmatrix} \quad (3)$$

$$R_c = \left[Y_W \frac{D}{R_W} \right] + (1 - D) R \quad (4)$$

The first stage is to transform the XYZ tristimulus values into sharpened cone responses using Equations 2 and 3. The linear von Kries transform with an incomplete adaptation term, D, is shown in Equation 4 for a single sensor. The primary difference between the iCAM chromatic adaptation transform and the CIECAM02 transform is in the definition of the whitepoint, R_w in Equation 4. The iCAM transform uses a low-pass version of the image itself as the adapting whitepoint to perform a localized adaptation. This adaptation can be a chromatic adaptation, as describe in Equations 2-4, or can be a luminance only adaptation. This can be accomplished by replacing Equation 2 with Equation 5 below.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \mathbf{M}_{CAT02} \begin{bmatrix} Y \\ Y \\ Y \end{bmatrix} \quad (5)$$

There are essentially three free parameters for the chromatic adaptation transform: the amount of blurring in the low-pass image, the degree of incomplete adaptation, and the choice between chromatic and luminance adaptation. In essence, however, each of these parameters depends upon the other choices. For instance, the degree of adaptation depends highly upon the choice of the blurring function. We utilize a Gaussian blur in the frequency domain, as shown in Equation 6.

$$Filter = e^{-\frac{x+y}{\sigma}^2}$$
$$IM_{filt} = FFT^{-1}\left\{ FFT\{\mathrm{Im}\} FFT\{Filter\} \right\} \quad (6)$$

The width of the filter, σ, is proportional to the amount of blurring. There are several options for choosing the amount of blurring. The first is to completely characterize the viewing conditions, and specify the width in device independent coordinates such as cycles-per-degree of visual angle. This approach is the ideal situation, but requires *a-priori* knowledge of the output viewing conditions. A simplifying assumption can be to specify the width of the filter in coordinates of the image itself. This suggests that the observer will alter the viewing conditions (by moving closer or farther) to get constantly sized images. Thus we can specify the σ in Equation 6 as a function of image size, in pixels such as xPixels/4.

The size of the blurring function will have a direct effect on the remaining free parameters. For instance, as we approach a large blurring function, a quarter or half of the image size, then the need for partial adaptation and luminance adaptation decreases. That is because the whitepoint image is converging upon the average of the entire scene. In the limit the blurring function becomes the average of the scene and the transform becomes identical to CIECAM02. In this situation, the partial adaptation function from CIECAM02 can be used. As the blurring function becomes smaller, then the need to specify the degree of partial adaptation explicitly becomes necessary. Figure 2 shows an example of this.

Figure 2. Simultaneous variation of blurring filter (columns) and degree of partial adaptation (rows).

Each column in Figure 2 represents a series of images filtered with two sized blurring kernels. The left column uses a kernel 1/4 the size of the image, while the right column uses a kernel 1/16 the size. The rows represent various degrees of partial adaptation (0.1, 0.5, and 1.0). The top row of images in Figure 2 look similar to each other, with a bit more shadow information in the left image, and more highlight information in the right image. As the degree of adaptation increases the difference between to two filter sizes become very evident. In both columns the overall color balance shifts from orange to "gray." In the extreme, if each pixel were adapted to itself then the entire image would go gray. This is fairly evident in the bottom right image. There is little color information remaining, and there are distinct "rings" around the bright light sources. This is a common phenomena found in many HDR tone rendering algorithms.[8,17]

From Figure 2 it is possible to infer some default settings for the parameters of iCAM for high dynamic range tone rendering. A filter size of 1/4 the image size seems to work well for most images, though again it should be stressed that this is based on an assumption for the viewing conditions. For unusually large or small images this parameter might need to be altered. The partial adaptation factor, D, should be set between 0.1 and 0.4 for most applications. If auto color balancing is desired, then the standard von Kries chromatic adaptation transform can be used. If the overall color balance needs to remain the same, replacing Equation 2 with Equation 5 results in just a luminance adaptation transform.

Local Contrast/Surround Effect

The chromatic and luminance adaptation only takes care of part of the HDR tone mapping. Once the HDR image has passed through the chromatic adaptation transform, the image is manipulated through a series of local tone reproduction curves. These curves can be thought of as changes in local contrast in the image, as a result of changes in the localized surround and luminance.

It is well understood in the field of color appearance that the overall luminance and surround effect perceived contrast of an image. In traditional color appearance there is a single surround and luminance factor (Dim, Dark, Normal, etc). In the iCAM image appearance a spatially localized "surround" map controls model the surround and luminance factor. This map is another low-passed version of the absolute Y image. This low-passed image is used to calculate a series of power functions, or tone reproductions curves. These functions are calculated by using the surround and luminance calculations from CIECAM02.[15] This is similar to the local contrast adjustments described by Moroney.[16]

$$F_L = 0.2\left[\frac{1}{(5L_A+1)}\right]^4(5L_A) + 0.1\left[1-\left[\frac{1}{(5L_A+1)}\right]^4\right]^2(5L_A)^{\frac{1}{3}} \quad (7)$$

The F_L factor from CIECAM02 is described in Equation 7., and shown in Figure 3 it is scaled by an additional 1.0/1.7 to normalize the function to 1 at a luminance of 1000 cd/m^2.

For use in iCAM the adapting luminance, L_A, is considered to be the low-passed version of the absolute Y tristimulus image.

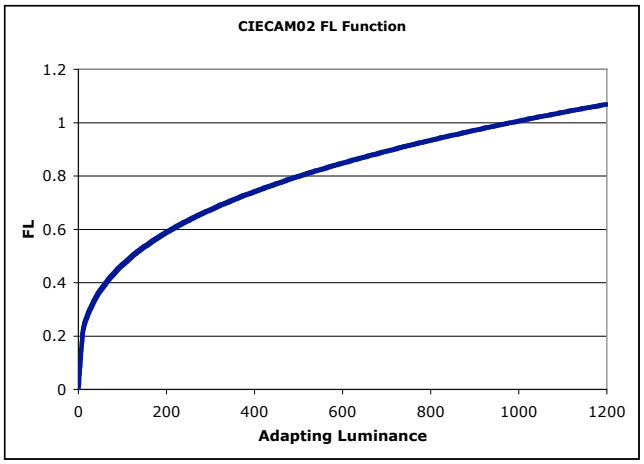

Figure 3. F_L surround luminance function from CIECAM02

The result is a tone curve for each pixel in the image, depending on the luminance of the surround neighborhood. This results in two free parameters: the normalization of the F_L function, and the width of the blurring kernel.

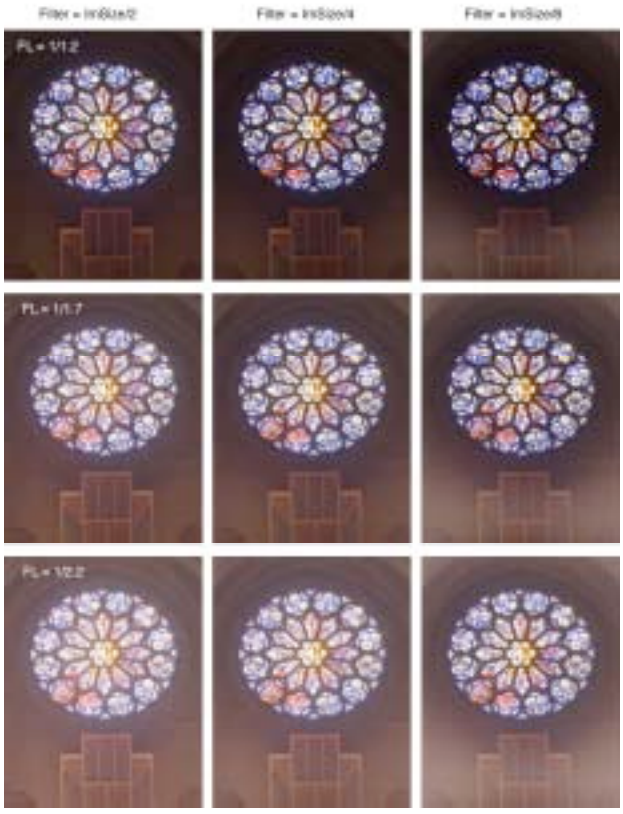

Figure 4. Effect of blurring kernel size, and F_L scaling factor

Figure 4 shows the effect of altering the blurring kernel size, using Equation 6 again, as well as changing the scaling factor of the CIECAM02 F_L curve. Essentially, changing the scaling factor "shifts" the image location around on the curve. As evident in Figure 4, increasing the scaling factor has an effect of increasing information in the shadow region, while desaturating the bright region. This is illustrated by moving down the rows in Figure 4. Likewise, increasing the size of the blurring kernel increases the information in the shadow areas, although at the cost of decreased saturation and increased "halo" effect around light sources. We have found the ideal setting is between 1/1.5 and 1/1.7 for the scaling factor, and a blurring kernel between 1/2 and 1/4 of the image size.

IPT Transform

The surround exponents calculated in the previous section are actually used in the transform from XYZ tristimulus values into the IPT appearance space. The first stage in this transform in to convert the XYZ units into LMS cone responses, as shown in Equation 8.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0 & 0.0 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix} \quad (8)$$

These cone responses are then compressed using a nonlinear power function, traditionally as shown in Equation 9 for the L channel.

$$\begin{aligned} L' &= L^{0.43}; \quad L \geq 0 \\ L' &= -|L|^{0.43}; \quad L < 0 \end{aligned} \quad (9)$$

This nonlinear power function is modified on a *per-pixel-basis* by the surround map calculated in the previous section, as shown in Equation 10.

$$\begin{aligned} L' &= L^{0.43 \cdot F_L}; \quad L \geq 0 \\ L' &= -|L|^{0.43 \cdot F_L}; \quad L < 0 \end{aligned} \quad (10)$$

Equation 10 is repeated for the M and S channels. Typically the LMS responses are brought into the IPT appearance space for calculation of appearance correlates such as lightness, chroma, and hue. This would be accomplished using Equation 11.

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \quad (11)$$

If we are not interested in the actual appearance correlates, and are simply interested in displaying the HDR tone mapped then this step is not necessary. To invert the IPT image back for display we can invert Equation 10 for a single surround condition. This is shown in Equation 12.

$$L' = L^{\frac{1}{0.43}}; \quad L \geq 0$$
$$L' = -|L|^{\frac{1}{0.43}}; \quad L < 0 \qquad (12)$$

The LMS cone responses, after applying the surround tone reproduction functions, are then converted back into CIE XYZ tristimulus values.

Displaying the Images

To display the XYZ images on a monitor we must first invert the chromatic adaptation transform, from D65 to the monitor whitepoint. This is accomplished using Equations 2-4 once again. The transformed XYZ values are then converted back to RGB using the inverse of Equation 1. The results are linear RGB values. The final images can be displayed by accounting for the display nonlinearity and scaling the images between 0-255. Often it is beneficial to apply a clipping function to the linear RGB data before scaling. This clipping function can remove any extremely bright pixels prior to display. The clipping is defined as a function of the image data itself, often as a percentile. For instance clipping to the 99th percentile of the image data, results in the following equation.

$$clip: RGB < 99\% = RGB$$
$$RGB > 99\% = 99\% \qquad (13)$$

Figure 5. Effects of clipping the RGB image prior to display

An example of clipping to various degrees is shown in Figure 5. The top frame illustrates no clipping. The highlight detail outside the parking garage is clearly visible, though the shadow detail is mostly lost. The middle frame shows clipping at 99%. There is information in both the outside highlights and inside shadow detail. The final frame shows clipping at 95%. The highlights are partly lost, but there is great detail in the shadow regions. One can pick the clipping level depending on the desired effect. For generally purpose tone-mapping a choice of 99% seems to work well.

After clipping the RGB values are then scaled so that the minimum value is 0 and the maximum is 1.0

$$RGB = \frac{RGB - RGB_{\min}}{RGB_{\max} - RGB_{\min}} \qquad (14)$$

Finally these values are compressed with a power function, and scaled between 0 and 255.

$$dRGB = 255 \cdot RGB^{\frac{1}{1.7}} \qquad (15)$$

The dRGB values can then be displayed normally on a monitor, or saved to an image file.

Conclusions

This paper has outlined the use of an image appearance model, iCAM, for tone mapping high dynamic range images. High dynamic range images are images that have a large contrast ratio and can exceed 5 units of magnitude. Tone mapping is necessary to display these images on a device with a much smaller dynamic range.

iCAM is a model designed for predicting overall image appearance rather than a specific tone-mapping algorithm. As such, it is uniquely suited for rendering high-dynamic range images in a perceptually meaningful way. This paper has outline the general use of the iCAM framework, and provided some specific implementation details that provide visually pleasing renderings. The parameters given in this paper can be thought of as a starting point for research on the perception of high dynamic range scenes. It is not the intent to imply that the parameters given here adequately describe the overall perceptual appearance. Extensive psychophysical research must be undertaken in order to make that claim. It is hoped that in the future iCAM will provide a foundation upon which to base that type of experimentation.

The general source code (Matlab, Mathematica, IDL) to iCAM, and this implementation for HDR tone mapping can be found at http://www.cis.rit.edu/mcsl/iCAM

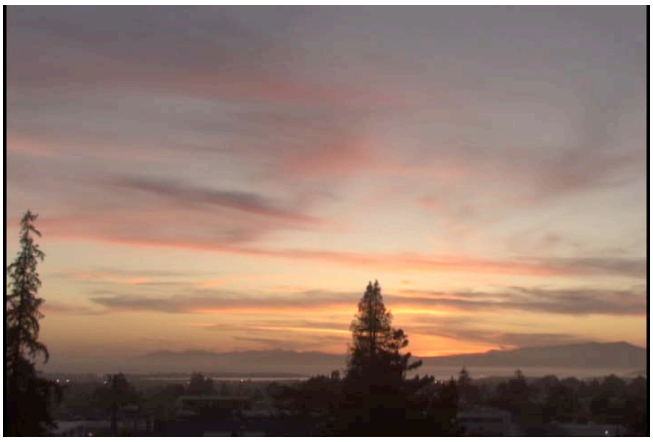

Figure 6. "Pleasing" renderings of HDR images from http://www.debevec.org

References

1 P. E. Debevec and J. Malik, Recovering High Dynamic Range Radiance Maps from Photographs, *Proc. SIGGRAPH'97*, 369-378 (1997).

2 F.Xiao, J.M. DiCarlo, P.B. Catrysse and B.A. Wandell High Dynamic Range Imaging of Natural Scenes, *IS&T/SID 10th Color Imaging Conference,* 337-342 (2002).

3 D. Yang, B. Fowler, A. El Gamal, and H. Tian, A 640x512 CMOS Image Sensor with Ultrawide Dynamic Range Floating-Point Pixel-Level ADC, *IEEE Journal of Solid State Circuits,* **34**, 1821-1834, (1999).

4 S. K. Nayar and T. Mitsunaga, High Dynamic Range Imaging: Spatially Varying Pixel Exposures, *Proc. IEEE CVPR*, (2000).

5 K. Takemura, Challenge for improving image quality of a digital still camera, *Proc. SPIE Electronic Imaging Conf. Santa Clara*, (2003).

6 G.Ward, High Dynamic Range Imaging, *Proc. IS&T/SID 9th Color Imaging Conference,* 9-16 (2001).

7 G. W. Larson, H. Rushmeier, and C. Piatko, A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes, *IEEE Transactions on Visualization and Computer Graphics*, **3**, 291-306, (1997).

8 E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, Photographic Tone Reproduction for Digital Images, *Proceedings of SIGGRAPH 02*, 267-276 (2002).

9 F.Durand and J. Dorsey, Fast Bilateral Filtering for the Display of High-Dynamic Range Images, *Proceedings of SIGGRAPH 02*, 257-266 (2002).

10 J. M. Dicarlo and B.A. Wandell, Rendering High Dynamic Range Images, *Proceedings of SPIE Image Sensors,* 392-401 (1999).

11 S.N. Pattanaik, J.A. Ferwerda, M.D. Fairchild, and D.P. Greenberg, "A Multiscale Model of Adaptation and Spatial Vision for Image Display," *Proceedings of SIGGRAPH 98,* 287-298 (1998).

12 S.N. Pattanaik, M.D. Fairchild, J.A. Ferwerda, and D.P. Greenberg, "Multiscale Model of Adaptation, Spatial Vision, and Color Appearance," *IS&T/SID 6th Color Imaging Conference,* Scottsdale, 2-7 (1998).

13 M.D. Fairchild and G.M. Johnson, Meet iCAM: An image color appearance model, *IS&T/SID 10th Color Imaging Conference,* 33-38 (2002).

14 F. Ebner, and M.D. Fairchild, "Development and Testing of a Color Space (IPT) with Improved Hue Uniformity," *IS&T/SID 6th Color Imaging Conference,* Scottsdale, 8-13 (1998).

15 N. Moroney, M.D. Fairchild, R.W.G. Hunt, C.J. Li, M.R. Luo, and T. Newman, "The CIECAM02 Color Apperance Model," *IS&T/SID 10th Color Imaging Conference*, Scottsdale, 23-27 (2002).

16 N. Moroney, "Local Color Correction Using Non-Linear Masking," *IS&T/SID 8th Color Imaging Conference*, Scottsdale, 108-111 (2000).

17 S. Pattanaik and H. Yee, Adaptive Gain Control for High Dynaimc Range Image Display, *Proceedings of 18th ACM Spring Confernce on Computer Graphics*, 83-87 (2002).

Rendering HDR Images

Garrett M. Johnson and Mark D. Fairchild
Munsell Color Science Laboratory, Center for Imaging Science
Rochester Institute of Technology, Rochester NY

Abstract

Color imaging systems are continuously improving, and have now improved to the point of capturing high dynamic range scenes. Unfortunately most commercially available color display devices, such as CRTs and LCDs, are limited in their dynamic range. It is necessary to tone-map, or render, the high dynamic range images in order to display them onto a lower dynamic range device. This paper describes the use of an image appearance model, iCAM, to render high dynamic range images for display. Image appearance models have greater flexibility over dedicated tone-scaling algorithms as they are designed to predict how images perceptually appear, and not for the singular purpose of rendering. In this paper we discuss the use of an image appearance framework, and describe specific implementation details for using that framework to render high dynamic range images.

Keywords

Color Appearance Models, Image Appearance Models, High Dynamic Range Imaging, Image Reproduction