A Comparison of Techniques for Rendering Scenes

Nathan Moroney and Ingeborg Tastl
Hewlett-Packard Laboratories, Palo Alto, CA

Abstract

Scene rendering or the conversion of higher dynamic range input generated by a capture device or synthetically to a lower dynamic range for reproduction on a different device presents unique processing challenges. This paper summarizes and compares several techniques for scene rendering. The paper specifically addresses the issue of dynamic range mapping, independent of demosaicing, camera calibration, flare compensation, and white point estimation. There are a number of computational, image attribute, application and visual phenomena to be considered. Qualitative comparisons are also provided for several of the algorithms.

Introduction

The challenges and difficulties of scene rendering have long been appreciated. There are often significant differences in the illumination and dynamic range of an original scene in comparison to a given reproduction. The complexities of the human visual system can further complicate the creation and assessment of a scene rendering. In *The Art of Light and Shadow*, Athanasius Kircher noted[1] that "our Roman Academy painters make errors when they attempt to paint daytime scenes by candlelight at night, except if by great and singular effort they can compensate for the crude shadows cast." Likewise, a digital camera or capture device must embody a method to convert the radiances captured on an area sensor to an acceptable color reproduction. This processing must take into account the fact that the luminance dynamic range of scenes can vary over several orders of magnitude, such as lows of 50:1 to highs of 50,000:1.

The earliest scene rendering was embodied in the aesthetic decisions made by artists during the creation of a painting or other work of art. Reverse engineering these decisions is a complex and elusive task. However even with the earliest crude photometric devices it was understood even in the nineteenth century that artists were not simply reproducing the photometry of the scene. With the invention and refinement of photography scene rendering was embodied in the design of film sensitivities and the associated photofinishing process that created prints from the negatives captured for a given scene. Much effort went into the optimization of a fixed set of tone curves that balanced quality and robustness of the scene rendering for a wide range of scenes.[2]

Consumer video cameras face similar problems[3] with rendering scenes and have severe computational limitations but have a temporal axis with which to adaptively modify scene rendering in real time. Other fields, such as computer graphics,[4,5] medical imaging,[6,7] illumination engineering,[8] machine vision,[9] and defense industries,[10,11] have also made use of increasingly sophisticated algorithms for processing, enhancing and analyzing high-bit depth or low contrast images. These other fields often have additional considerations and approaches to solving the optimization of quality of rendering with robustness.

There has also been focused and thorough consideration of scene rendering for digital photography.[12] In this case rendering refers to the conversion of scene-referred data to output-referred data, such as creating an sRGB image from sensor data. This rendering can also include gamut mapping or preference adjustments. It is also possible to convert from one scene-referred image to another scene-referred image, such as re-lighting a scene after capture or scene-editing. Finally, re-rendering or repurposing is the conversion from one output referred encoding to another output referred encoding, such as converting an sRGB image to the ICC PCS. It should be noted that the use of tone curves in digital photography achieves more than dynamic range compression. For instance tone curves have been used for changes in viewing conditions, chroma boosts and preference adjustments.

Parallel to the advance and diversification of technologies incorporating scene rendering techniques or spatial image enhancements was continued progress in understanding and modeling the human visual system.[13-18] The CIE with basic colorimetry and appearance modeling provides a set of tools and terminology for describing colors in aperture mode or other relatively simple viewing conditions. The Retinex model provided an alternative approach that emphasized and modeled the importance of spatial factors. There are a large number of related algorithms refining the basic model but there has been recent progress towards a reference implementation.[19] Other researchers have developed vision models which incorporate spatial aspects of perception using multi-scale representations of a scene, such as difference of gaussians or oriented difference of gaussians. These vision models are valuable in that they provide algorithms and techniques that have been or can be applied to scene rendering.

Specific Techniques

Many early techniques made use of a tone reproduction curve or tone curve to map input values to some new output value. Often these tone curves were fixed for a device or at a system level. For example, the s-shape or sigmoidal curve is used in many film systems. These curves have evolved and recent curves tend to focus on highlight compression versus shadow compression.[20] Likewise, early video cameras or studio cameras used a fixed tone curve or gamma when capturing a scene.

A significant advance is the image-specific tone curve. The technique described by Holm[21] renders images based on the properties of both the input image and the dynamic ranges of the input and output device. This approach was optimized for digital photography and was embodied as red, green and blue tone curves. Braun and Fairchild[22] proposed image specific device to device reproduction using sigmoidal transfer functions applied to the L* or lightness channel of images. This technique could be described as contrast reproduction for previously rendered images. An important benchmark for image specific tone curves is manual rendering or the visual inspection and optimization of a tone curve by a human observer. Manual rendering, while impractical at a consumer level, is informative in that it provides information about observer preferences that none of the current vision models address.

Pattanaik et al.[13] proposed a visual multiscale visual model for scene rendering. This model was applied to both computer graphics and actual scenes. It makes use of a six level difference-of-Gaussian pyramid and uses a two-part power function for the adapted contrast transducers. The Reihard et al. model[20] also use a series of normalized difference-of-gaussian images. Their model was derived for photographic rendering and computer graphics and uses a local contrast measure to achieve dynamic range compression.

The retinex model has a long history and a large number of derivative algorithms.[23] The basic algorithm makes use of edge ratios, a reset operator and a specific path through an image. The radiance ratio operation is often implemented as differences of log radiances for efficiency. The algorithm is both surprisingly simple and subtlety complex. Variations on the algorithm include a ratio modification operator, a pyramidal implementation, an interpretation within a variational framework and a wavelet implementation. Recently McCann has demonstrated the importance of low spatial frequencies for understanding visual perception.[15]

A last category to consider is manual dodging and burning of images. Like manual tone curves, this technique is impractical for consumer imaging, but it is informative as an ideal benchmark for any of the spatially operating algorithms. Furthermore, it can reveal preference judgments that may actually contradict vision models. For example for compositional purposes it may be better to reduce contrast of a complex region in an image that is competing with the subject of the image for visual attention.

Additional Considerations

There are many visual phenomena to consider with respect to scene rendering. The perception of lightness and brightness for simple visual stimuli is both well modeled and the subject of ongoing research. A non-linear compressive function is often used to relate the luminance of a stimulus and a white point to its lightness or brightness. However, simultaneous contrast effects must also be considered. In this case the lightness of the area immediately surrounding the stimulus will affect the perceived lightness. This has been modeled by a variable background-dependent exponent. Crispening or the apparent increase in perceived differences in stimuli relative to a background is another phenomena to consider. This has been modeled using a variable sigmoidal function. Other phenomena which may be related include chromatic adaptation transforms, adopted white point, adopted white luminance, multiple white points, depth perception, transparency, anchoring, media and even perception of chroma. Ongoing research in retinal "wiring" and cortical processing also have significant implications for vision models, and by extension scene rendering algorithms that attempt to emulate the visual system.

Scene rendering techniques create a color reproduction of scene and directly impact a number of specific properties or features. For example these algorithms will establish the white and black points of the images. Likewise, they will establish the overall contrast and brightness of the image and will also balance global and local contrast if they are spatial. These algorithms will also likely impact gray balance, chroma reproduction and detail visibility.

Scene rendering, like other image processing operations, has associated benefits and potential artifacts. One commonly noted artifact is haloing or the ringing of edges with a gradient. There is also the potential to magnify noise in addition actual image detail when rendering scenes. All techniques attempt to achieve an optimal contrast for the reproduction. In addition, spatial algorithms must balance global and local contrast. Depending on the input format or color space for the image, there is also the possibility for chroma or hue distortions. Finally, care must be taken that an image is not over-corrected or processed multiple times.

Examples

Figures 1 through 4 show four renderings of the same image. Figure 1 is the original as captured and rendered using a consumer digital camera. Figure 2 through 4 are alternative renderings of Figure 1. Figure 2 shows the original image re-rendered using a preferred photographic rendering.[21] Figure 3 shows the original image re-rendered using a modified Retinex algorithm. The primary modification is the use of Sobol's ratio modification operator.[24] Figure 4 shows the original image re-rendered using a modified mask-based rendering.[25]

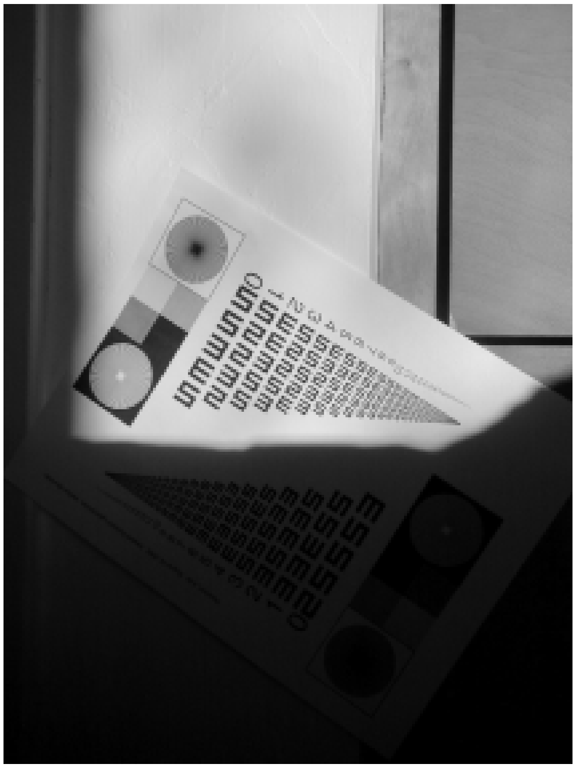

Fig. 1. Original digital photograph.

Fig. 3. A modified Retinex rendering.

Fig. 2. A preferred photographic rendering.

Fig 4. A modified mask-based rendering.

The original image in Figure 1 has a number of imperfections and was processed as an 8 bit image. Figure 2 shows that a carefully designed tone curve can considerably improve even difficult high dynamic range scenes. Figure 3 is considerably improved relative to the original and has minimal spatial artifacts. Figure 4 is similar to Figure 3 and shows that for some images, the two approaches yield similar results. The primary difference between Figure 3 and 4 is that the modified mask-based rendering allows highlight information to be darkened while the shadows are lightened.

Conclusion

This paper describes and compares a set of techniques for rendering scenes. Specific considerations, related perceptual phenomena and potential image attributes and artifacts are considered. Finally an example is provided for a qualitative comparison of the three rendering algorithms. Considerable improvement can be achieved with a carefully designed tone correction curve. However, spatial processing such as Retinex or mask-based processing provides additional improvement. Mask-based processing can darken and lighten image regions while Retinex only lightens image regions.

References

1. Athanasius Kircher, *Ars Magna Lucis et Umbrae*, Amsterdam, p. 142 (1671).
2. T. H. James. Ed., *The Theory of the Photographic Process*, Eastman Kodak Company, Rochester NY, 536-561 (1977).
3. J.Y. Kim, L.S. Kim and S.H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization", *IEEE Transactions on Circuits and Systems for Video Technology*, **11**(4), 475-484 (2001).
4. J. Tumblin and G. Turk, "LCIS: A Boundary Hierarchy for Detail Preserving Contrast Reduction", *SIGGRAPH 99*, 83-90 (1999).
5. Paul E. Debevec and Detendra Malik, "Recovering High Dynamic Range Radiance Data from Photographs", *SIGGRAPH '97,* 369-378 (1997).
6. S. Dippel, M. Stahl, R. Wiemker and T. Blaffert, "Multiscale contrast enhancement for radiographics: Laplacian pyramid versus fast wavelet transform", *IEEE Transactions on Medical Imaging*, **2**4(4) 343-353 (2002).
7. J. Yinpeng, L. Fayad and A.F. Laine, "Contrast enhancement by multiscale adaptive histogram equalization", *SPIE Proc. Wavelets: Applications in Signal and Image Processing IX*, **447**8, 206-213 (2001).
8. I. Ashdown, "Luminance gradients: photometric analysis and perceptual reproduction", *J. Illuminating Engineering Soc*. 25, 69-82 (1996).
9. C. Zikuan, T. Yao and C. Xin, "Multiresolution local contrast enhancement of X-ray images for poultry meat inspection", *Applied Optics*, **40**(8), 1195-1200 (2001).
10. John S. Dehne, "Image processing techniques for real time imagery", *Proceedings of the NATO Advanced Study Institute on Pattern Recognition and Signal Processing*, Sijthoff and Noordhoff, Netherlands, 205-224 (1978).
11. P.M. Narendra and R.C. Fitch, "Real-time adaptive contrast enhancement", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. **PAMI-3**, 6, (1981).
12. ISO/CD 22028-1, *Extended colour encodings for digital still image storage, manipulation and interchange – Part 1: Architecture and requirements* (2002).
13. S. N. Pattanaik, J.A. Ferwerda, M. D. Fairchild and Donald P. Greenberg, "A multiscale model of adaptation and spatial visual for realistic image display", *SIGGRAPH 98*, 287-298 (1998).
14. B. Blakeslee and M.E. McCourt, "A multiscale spatial filtering account of the Werthheimer-Benary effect and the corrugated Mondrain", *Vision Res.*, **41**, 2487-2502 (2001).
15. J. McCann, "Calculating lightness in a single depth plane", *J. Electronic Imaging*, **10**(1), 110-122 (2001).
16. Edwin H. Land and John McCann, "Lightness and the Retinex Theory", *J. Opt. Soc. Am.*, **61**, 1-11, (1971).
17. A.L. Gilchrst, *Lightness, Brightness and Transparency*, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ (1994).
18. E.H. Adelson, "Lightness perception and lightness illusions", *The New Cognitive Neurosciences*, MIT Press, Cambridge, MA, 339-351 (1999).
19. B. Funt, F. Ciurea, and J. McCann, "Retinex in Matlab", *Proc. 8th Color Imaging Conference*, 112-121 (2000).
20. E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic tone reproduction for digital images", *SIGGRAPH 02*, (2002).
21. J. Holm, "Photographic Tone and Colour Reproduction Goals," *Proc. CIE Expert Symposium 96: Colour Standards for Image Technology*, 51-56 (1996).
22. Gustav J. Braun and Mark D. Fairchild, "Image lightness rescaling using sigmoidal contrast enhancement function", *J. Elect. Imaging*, **8**, 380-393 (1999).
23. Various authors, Session Track: "Retinex at 40", *Human Vision and Electronic Imaging VII*, B.E. Rogowitz and T.N. Pappas Eds. *Proc. SPIE* **4662**, 278-439 (2002).
24. Robert Sobol, "Improving the Retinex algorithm for rendering wide dynamic range photographs", *Human Vision and Electronic Imaging VII*, B.E. Rogowitz and T.N. Pappas Eds. *Proc. SPIE* **4662**, 341-348 (2002).
25. N. Moroney, "Local color correction using non-linear masking", *Proc. 8th IS&T/SID Color Imaging Conference*, 108-111 (2000).

Biography

Nathan Moroney is a senior color researcher in the Color Imaging and Printing Technologies Department of Hewlett-Packard Laboratories. Previously, he worked for the Barcelona division of Hewlett-Packard and at the RIT Research Corporation. He has a Masters Degree in Color Science from the Munsell Color Science Laboratory of RIT and a Bachelors degree in color science from the Philadelphia University. His research interests include color appearance models and digital color imaging pipelines. He is a member of the IS&T and the ISCC. Nathan is currently the technical chair for CIE TC8-01.