

A Wide Field, High Dynamic Range, Stereographic Viewer

Greg Ward
Exponent – Failure Analysis Assoc.
Menlo Park, California

Abstract

We present a high dynamic range viewer based on the 120-degree field-of-view LEEP stereo optics used in the original NASA virtual reality systems. By combining these optics with an intense backlighting system (20 Kcd/m^2) and layered transparencies, we are able to reproduce the absolute luminance levels and full dynamic range of almost any visual environment. This technology may enable researchers to conduct controlled experiments in visual contrast, chromatic adaptation, and disability and discomfort glare without the usual limitations of dynamic range and field of view imposed by conventional CRT display systems. In this paper, we describe the basic system and techniques used to produce the transparency layers from a high dynamic range rendering or scene capture. We further present an empirical validation demonstrating device's ability to reproduce visual percepts, and compare this to results obtained using direct viewing and a visibility matching tone reproduction operator presented on a conventional CRT display.

Introduction

The natural world presents our visual system with a wide range of colors and intensities. A starlit night has an average luminance level of around $10^{-3} \text{ candelas/m}^2$, and daylight scenes are close to 10^5 cd/m^2 . Humans can see detail in regions that vary by $1:10^4$ at any given adaptation level, over which the eye gets swamped by stray light (i.e., disability glare) and details are lost. Modern camera lenses, even with their clean-room construction and coated optics, cannot rival human vision when it comes to low flare and absence of multiple paths (“sun dogs”) in harsh lighting environments. Even if they could, conventional negative film cannot capture much more range than this, and most digital image formats do not even come close. With the possible exception of cinema, there has been little push for achieving greater dynamic range in the image capture stage, because common displays and viewing environments limit the range of what can be presented to about two orders of magnitude between minimum and maximum luminance. A well-designed CRT monitor may do slightly better than this in a darkened room, but the maximum display luminance is only around 100 cd/m^2 , which doesn't begin to approach

daylight levels. A high-quality xenon film projector may get a few times brighter than this, but they are still two orders of magnitude away from the optimal light level for human acuity and color perception.

In this paper, we present a novel device for displaying stereographic, high dynamic-range images to a single viewer. This static display device combines simple, known technologies in a way that faithfully reproduces the visual field of a natural or simulated scene, resulting in a heightened sense of realism and presence. Following a description of the device and the requisite image processing techniques, we demonstrate the quality of its output by comparing visibility in the original scene to visibility in our HDR reproduction. We conclude with some ideas for future development of this technology in cooperation with Bristol University in England, and the University of British Columbia in Vancouver.

Device Description

The high dynamic-range viewer itself is a simple device, consisting of a bright, uniform backlight joined to a set of LEEP ARV-1 optics. The optics allow a 120° field of view in each eye for a very natural and complete stereo view from a pair of 2.5” (6.4 cm) fisheye images mounted side-by-side. A schematic of the viewer is shown in Fig. 1, and the actual viewer is shown in Fig. 2. Four viewers have been constructed thus far, and this is the latest model.

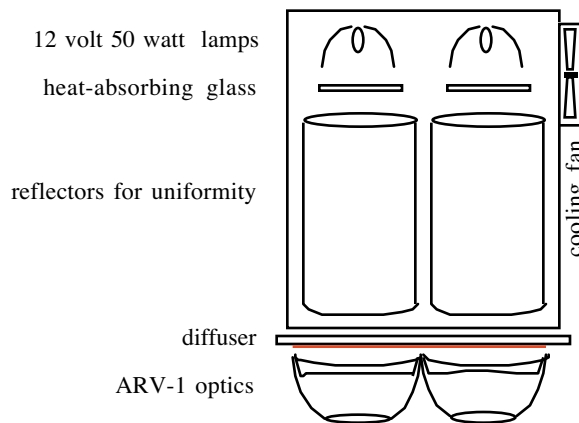


Figure 1. Schematic of HDR viewer.

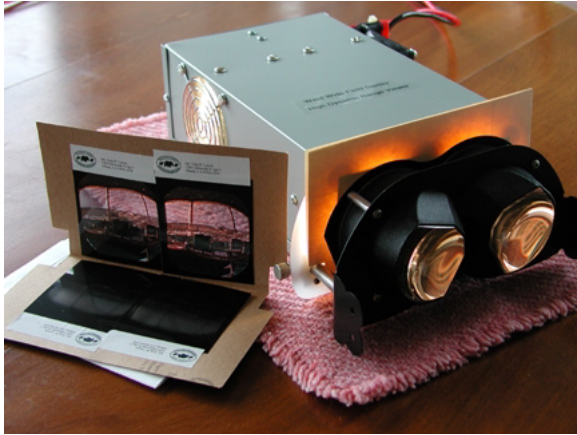


Figure 2. HDR viewer device.

The transparencies sit on top of the diffuser in front of the ARV-1 optics. Focus adjustment is performed by shifting the optics slightly closer or away from the transparencies, which are held in place by a clip. Precise focus is not important, as the eye will accommodate over a certain range. It is important when viewing to remove eyeglasses that might prevent the eyes from getting close enough to the optics to provide the full field of view. Provided the wearer does not have severe astigmatism, the focus adjustment permits acuity at the visual limit of 20 cycles/degree, provided the transparency is rendered at a density of 800 dpi or better (2048×2048 resolution).

Image Processing

The basic transformation required for correct perspective viewing with the ARV-1 optics is a hemispherical fisheye projection. This projection is most easily explained in the form of a diagram, as shown in Fig. 3.

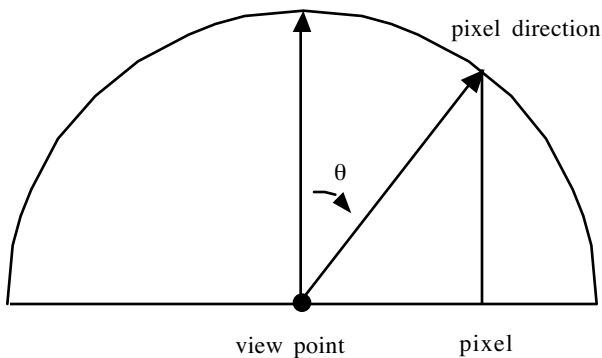


Figure 3. The hemispherical fisheye projection.

In a hemispherical fisheye projection, the distance from the center of the image is equal to the sine of the angle from the principal view axis (i.e., depth). This can be visualized as a hemisphere placed on the image plane and projected directly

down onto it, as shown in Fig. 3. Each pixel in the image corresponds to the ray that passes through the point in the hemisphere directly above it. In vector terms, this ray is defined by the origin view point and the unit direction vector given in Eq. 1. The x and y values in the equation are the offsets from the image center normalized to half the width of a full 180° image. These coordinates will equal 1 or -1 at the image edges, and 0 in the center. The unit view vectors for x , y , and z correspond respectively to the horizontal, vertical, and depth directions for this image. Note that the view is not defined if $(x^2 + y^2) > 1$.

$$\hat{v}_p = \hat{v}_x x + \hat{v}_y y + \hat{v}_z \sqrt{1 - x^2 - y^2} \quad (1)$$

Since each image covers only 120° rather than 180°, the only the corners of the image are perpendicular to the principal view axis. The final projection can be seen in the processed image shown in Fig. 4.

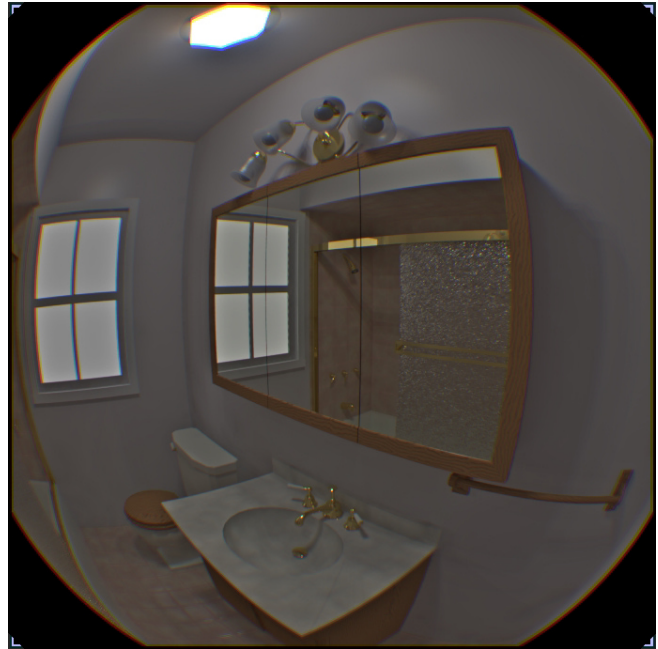


Figure 4. A detail transparency layer for the HDR viewer.

Due to chromatic aberration in the LEEP ARV-1 optics, it is best to precorrect the image by scaling the red image plane proportionally more than the blue image plane so that the red is about 1.5% larger than the blue, and the green is in between. This can be seen as a red fringe near the edges of Fig. 4, which results in fairly good color convergence when viewed through the LEEP optics. The middle of the image not subject to the same chromatic distortion, so the center of view is unaffected.

Due to limitations in the film recording process, it is not possible to achieve the full, desired dynamic range in a single transparency. Although the film recorder we used is capable of producing nearly 1000:1 at the limit, the bottom reaches of this range have fairly large intensity steps. The useful range where the intensity steps are below the visible

threshold necessary to avoid banding artifacts is closer to 100:1. Since we wish to produce images with a dynamic range in excess of 10,000:1, we need to use two transparencies, layered one on top of the other. By combining the ranges in this way, we double our dynamic range in log terms as the densities add together.

To avoid problems with transparency alignment and ghosting, we reduce the resolution of one layer using a Gaussian blur function. Because the dynamic range of the individual color channels is not important for perception, we convert to gray in the scaling layer to simplify our calculations. The resulting image is shown in Fig. 5. The degree of blur is not critical – we use a blur radius equivalent to 16 cycles across the image. We have found this to allow for easy image alignment without seriously compromising the maximum luminance gradient in the combined result.



Figure 5. A scaling transparency layer for the HDR viewer.

Since the combination of the two transparency layers is equivalent to adding the densities, each image must have half the density of the original. This is easily accomplished by taking the square root of each pixel value in the blurred image, and dividing the original image by this image for the detail layer. This result is the one shown in Fig. 4. The effects of dividing by the scaling layer may be seen in the odd gradients near the window's edge, which disappear when the layers are placed together in the HDR viewer. Alignment marks at the corners of the images aid in their registration.

So far, we have discussed only the generation of a single layered image, whereas two images are required for a stereo view, one for each eye. In cases where stereo perspective is negligible or unimportant, a single image may be duplicated for each eye to achieve a satisfactory monocular view. For stereoscopic viewing, one must capture or render two images, and this is done by simply

shifting the real or virtual camera by the average interocular distance, which is approximately 2.5" (6.4 cm). It is not necessary to adjust the principal view axis in the two perspectives, as the observer should and will make his or her own accommodation when viewing the stereo pair.

Validation Experiments

We have performed two types of validation for the HDR viewer's performance, one quantitative and one qualitative. In the quantitative study, we wanted to measure the luminances produced by the viewer and compared them to the original input to determine that the method of combining transparency layers performs as specified. In the qualitative study, we took several subjects into a darkened room and presented them with a contrast visibility chart, which we then reproduced in our HDR viewer. We start by discussing the quantitative study, and follow with a presentation of the qualitative results.

Quantitative Study

The quantitative measurement process presented some challenges, as there are no luminance probes with sufficient spatial resolution and freedom from stray light to measure the very large gradients we produce in the viewer. We therefore restricted our quantitative validation to a simple verification that combining transparencies adds densities as predicted. To accomplish this, we used an industry standard densitometer made by X-rite, and found that layering transparencies corresponds to adding densities within the 2% accuracy of the densitometer. Mathematically, the greatest theoretical error will happen near the lowest densities, where multiple interreflections could increase the effective transmittance by as much as 1%. This is not enough of an error to be significant in most simulations.



Figure 6. A contrast chart designed by Tom Ayers [Ayers96].

Qualitative Study

In our qualitative comparison, we asked subjects to view the chart shown in Fig. 6 in a darkened room, illuminated dimly by a single, distant spotlight designed to simulate a car's headlights. We then photographed the same chart using an Olympus E-10 digital camera and a method for combining multiple exposures into a calibrated, high dynamic-range image [Matsunaga&Nayer99]. This calibrated image is shown in Fig. 7.

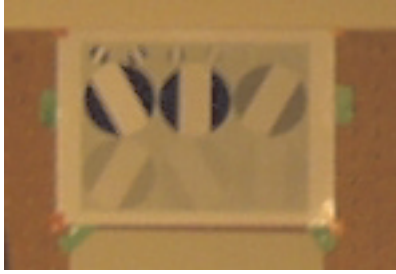


Figure 7. A high dynamic-range photograph of the contrast chart under the initial viewing conditions.

Unfortunately, the resolution of the photograph in Fig. 7 is not very high, and this turned out to be a problem for our tests. Although we could have used a longer focal length on our camera and thus captured better resolution in the chart image, we needed to capture as much of the surround as possible in order to have proper adaptation in the HDR viewer. The results for the smaller disk visibilities should not be taken too seriously for this reason.

In a second test, we shone a bright spotlight directly in the subjects' eyes directly under the chart to simulate a glare condition akin to a car's oncoming headlights and asked for which disks they could discern the orientation. This same condition was reproduced in the HDR viewer by introducing a white disk to the transparency with the appropriate size and position to produce the same effective luminous power on the retina as in the test condition.

The results of these two test conditions are shown for two subjects in Table 1 for the original scene and the HDR viewer. The viewing results give the number of large and small disks discernable, respectively. (In our tests, the visibility ordering of the disks never changes, so there is no need to account for the individual disk visibilities.)

Table 1. Disk Visibility: Real vs. HDR viewer.

| Subject | Condition | Real | HDR viewer |
|---------|-----------|------|------------|
| GW | normal | 5, 2 | 5, 0 |
| | glare | 2, 0 | 2, 0 |
| CS | normal | 6, 2 | 5, 0 |
| | glare | 3, 0 | 3, 0 |

As we can see in our results, the smaller disks were never quite visible in the HDR viewer, and we believe this is due to the limited resolution of the original photograph rather than a limitation of the viewer itself. We would like to repeat this experiment using a higher resolution image from a line-scan panoramic camera, but at the time of this

writing, we have not yet managed to do so. On the other hand, the large disk visibility was reproduced fairly well in our experiment, indicating that the basic technique of recreating luminances in a wide-field stereo viewer is an effective means of virtual reenactment in a luminance range corresponding to roadway illumination conditions.

One problem we noticed with reproducing the glare condition was that the interreflections in the HDR viewer optics were a source of difficulty. In particular, the bright region corresponding to the oncoming headlight reappeared as a ghost in the view, obscuring other parts of the image in the same way that "sun dogs" appear in a photograph containing the solar disk. One way to alleviate this problem is to coat the optics to reduce reflections, and this may be a desirable enhancement to future versions of the viewer. It is also helpful to use the viewer in a darkened room if the scene being reproduced is very dim, as stray light can otherwise enter from the sides and obscure the view and adversely affect viewer adaptation.

For comparison purposes, Fig. 8 shows the scene as simulated by the visibility-matching tone reproduction operator of Larson et al. [Larson97]. Fig. 8a shows the chart under the "normal" condition, and Fig. 8b shows the chart under the glare condition. Unless printing is carefully controlled, there is no guarantee that the print you see in the proceedings will reproduce the target visibility, but it worked reasonably well on our monitor.



Figure 8a. A visibility-matching tone operator applied to the normal chart viewing condition.

Conclusions

We have presented a simple device for displaying high dynamic-range, wide-field stereo imagery. The most serious limitation of the current device is its restriction to static scenes. Clearly, it would be better if we could quickly change from one image to another, without needing to swap transparencies and interrupt the observer. Ideally, we would

like to present animations of virtual reenactments or reproductions at video frame rates. The challenge of achieving the necessary resolution and bandwidth, although significant, may be met with today's PC graphics hardware. It may require multiple cards in the same machine and specialized drivers, but it is possible.

A group of researchers in the physics department at the University of British Columbia in Vancouver under the direction of Lorne Whitehead have already constructed a prototype high dynamic-range display with animation capabilities. Although their design is currently restricted to a typical monitor's field of view, there is good reason to believe that the same technology could also be applied in a wide-field stereo configuration. Another research group in the computer science department of Bristol University in England is interested in doing just that, and the author hopes to collaborate with both groups to achieve these goals.

References

1. Tom Ayers, Psychophysical Validation of Photographic Representations, *ASME* 1996.

2. Greg Ward Larson, Holly Rushmeier, Christine Piatko, A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 3, No. 4, December 1997.
3. T. Mitsunaga and S. K. Nayar, Radiometric Self Calibration, *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, June 1999.

Biography

Greg Ward (a.k.a. Greg Ward Larson) graduated in Physics from UC Berkeley in 1983 and earned a Master's in Computer Science from SF State University in 1985. Since 1985, he has worked in the field of light measurement, simulation, and rendering variously at the Berkeley National Lab, EPFL Switzerland, Silicon Graphics Inc., Shutterfly, and Exponent. He is author of the widely used *Radiance* package for lighting simulation and rendering.