

High Dynamic Range Imaging Pipeline: Perception-Motivated Representation of Visual Content

Rafał Mantiuk^a and Grzegorz Krawczyk^a and Radosław Mantiuk^b and Hans-Peter Seidel^a

^aMPI Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany;

^bSzczecin University of Technology, Żołnierska 49, 71-210 Szczecin, Poland

ABSTRACT

The advances in high dynamic range (HDR) imaging, especially in the display and camera technology, have a significant impact on the existing imaging systems. The assumptions of the traditional low-dynamic range imaging, designed for paper print as a major output medium, are ill suited for the range of visual material that is shown on modern displays. For example, the common assumption that the brightest color in an image is white can be hardly justified for high contrast LCD displays, not to mention next generation HDR displays, that can easily create bright highlights and the impression of self-luminous colors. We argue that high dynamic range representation can encode images regardless of the technology used to create and display them, with the accuracy that is only constrained by the limitations of the human eye and not a particular output medium. To facilitate the research on high dynamic range imaging, we have created a software package (<http://pfstools.sourceforge.net/>) capable of handling HDR data on all stages of image and video processing. The software package is available as open source under the General Public License and includes solutions for high quality image acquisition from multiple exposures, a range of tone mapping algorithms and a visual difference predictor for HDR images. Examples of shell scripts demonstrate how the software can be used for processing single images as well as video sequences.

Keywords: high dynamic range, image representation, software, pfstools, computer graphics

1. INTRODUCTION

A common image formats, such as JPEG or GIF, although enormously popular, are very limited in terms of preserving dynamic range and color gamut of real-world scenes. To address this problem, several custom image formats, tailored to the requirements of a particular application and discipline, have been developed. These include high dynamic range (HDR) formats used in computer graphics (RGBE, OpenEXR); camera RAW formats used in photography; the DICOM image format used in medical imaging; or the DPX format used in cinematography. These formats, although superior to standard images in terms of fidelity, cannot gain a wide-spread acceptance because of lack of standards and efficient compression algorithms. But, before such standards can be established and new compression algorithms are developed, there is a need for a high-level view on the general imaging pipeline, its goals and limitations.

High dynamic range imaging (HDRI) is a new discipline that involves a highly accurate representation of images. As it originates from light simulation (computer graphics rendering) and measurements, the pixels of HDR images are assigned a physical meaning, while most imaging disciplines are more relaxed about what information a pixel contains. This highly accurate representation of images gives an unique opportunity to create a common imaging framework, that could meet the requirements of different imaging disciplines. HDRI could also foster a unified terminology and better understanding of imaging from the perspective of both the physics of light and visual perception.

In this paper we present a holistic approach to the imaging pipeline (Section 2) and disentangle some imaging terminology (Section 3). The concept of high dynamic range images and its advantages are discussed in Section 4. We then describe a software package that operates on HDR images (Section 5). The software has been developed for research purposes and made available as an open source project. The application of this software is shown on the examples of HDR image acquisition (Section 6), tone mapping (Section 7), and fidelity/quality assessment (Section 8).

Further author information: (Send correspondence to R.M.)

R.M.: E-mail: mantiuk@mpi-sb.mpg.de, Telephone: +49 681 9325-427

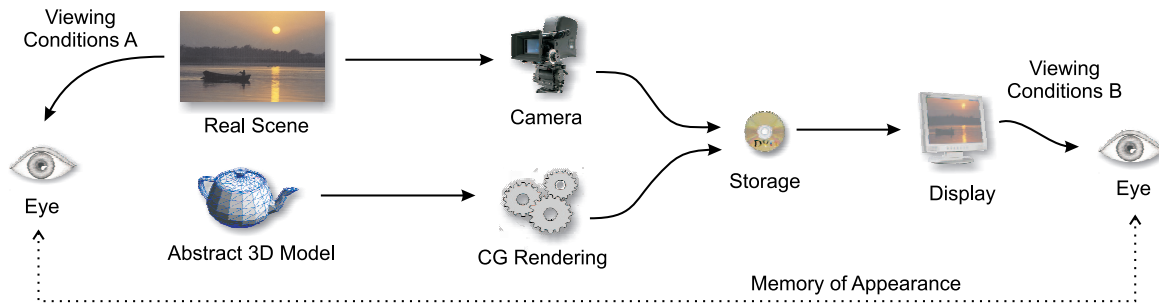


Figure 1. Example of an imaging pipeline — from acquisition to display. Different stages of processing involve different representation of an image.

2. IMAGING PIPELINE

Figure 1 illustrates an example of an imaging pipeline that starts with acquisition of a real world scene or rendering of an abstract model using computer graphics techniques and ends at a display. Each stage of a typical pipeline operates on images, which in theory should be 2D arrays of light measurements, but in practice they are designed to evoke the most realistic impression of brightness, hue and colorfulness of the original scene with the rigid assumption of a display device. The processing associated with each stage usually aims at delivering the best looking images, which are probably sharper and more colorful than the real scenes. The representation of an image at each stage of the imaging pipeline often depends on the technology used, for instance, on the “gamma correction” of a camera or a display (more on this in the next section), on different color characteristics of a device and the algorithms that have been used to “beautify” an image. As the result, each stage of such a pipeline operates on very vaguely defined data, which is difficult to standardize (for example, many years of effort have been needed to standardize the ICC profiles).

The only reason why the actual imaging pipelines do an acceptable job is because the human eye has a huge tolerance for all imperfections that are present in images. Prints or images displayed on monitors are almost never compared with real scenes, but rather with their memorized appearance, which is imprecise.¹ Also, the difference in viewing conditions (illumination, field of view, etc.) between the reality and the reproduction does not allow to deliver precisely the same visual stimuli. Therefore, the goal of most imaging pipelines is to reproduce the appearance of a scene, but not necessary to deliver the same physical stimuli to our eyes.

The reason why the existing imaging pipelines start to reveal their shortcomings is a technical progress in both digital cameras and displays. There are cameras on the market that can capture a scene of over 120dB (e.g. HDRC, OMRON Stuttgart), compared to about 55dB in the case of traditional cameras. There are also HDR displays that can output the contrast of 1 : 50 000,² compared to 1 : 2 000 for the latest LCD displays. However, an image taken with a traditional camera can not take the full advantage of an HDR display because of the limited dynamic range of the photograph. Similarly, the output from an HDR camera is heavily distorted when stored using existing image formats and cannot be directly displayed on traditional displays. The stages of an imaging pipeline become incompatible if the devices of higher dynamic range or color gamut are used. Therefore, the imaging pipelines must be thoroughly redesigned in order to take advantage of these new technologies. In the next sections we explain several quantities associated with HDR representation of images and then we outline the concept of HDR imaging, which addresses the shortcomings of existing solutions.

3. TERMINOLOGY

In this section we explain several physical and perceptual quantities important for digital imaging. We do not give a complete or exhaustive introduction to radiometry, photometry or colorimetry, since these are described in full extend elsewhere,^{3,4} but we rather focus on the concepts that are fundamental to HDR imaging.

3.1. Light

As any complex phenomena, light can be described by multitude of quantities. The most accurate, in physical sense, is a radiometric measure of **spectral radiance**. Spectral radiance is a measure of photon flux of particular wavelength traveling

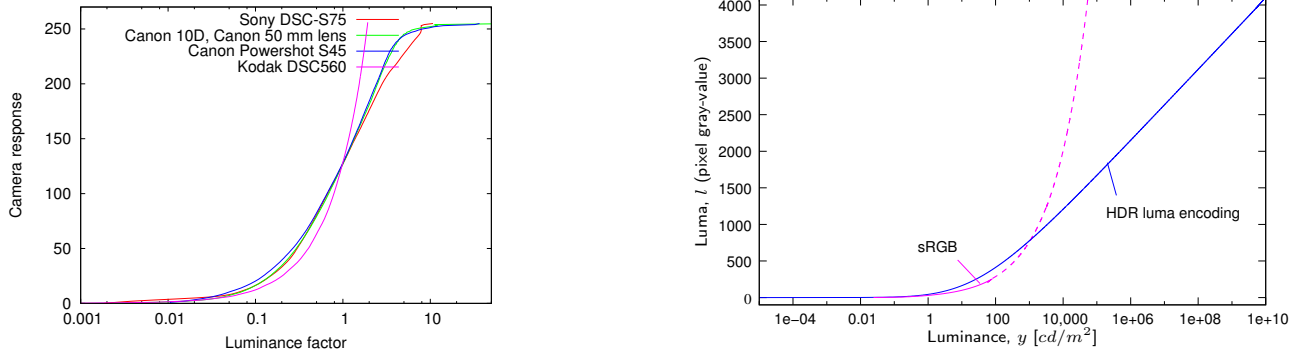


Figure 2. Left: Transfer functions of several digital cameras. The responses of red, green and blue color components are averaged and represented as a single line. The curves were measured using *pfscalibration* software described in Section 6. **Right:** Luminance to luma mappings derived from the characteristic of the HVS⁵ (blue) compared with the non-linearity of the sRGB color space (pink). The power function used in sRGB can not be used for the full visible range of luminance.

through an infinitively small area in particular direction (solid angle) at a unit of time. Spectral radiance is commonly used in light transfer simulations, such as computer graphics rendering algorithms that solve global illumination problem. Since spectral radiance needs to be specified for a continuum of wavelengths, it would be a cumbersome measure in many applications. Therefore, photometry defines a measure of **luminance**, which integrates spectral radiance over all visible wavelengths using the luminous efficiency function V . The function V approximates sensitivity of the human visual system to spectral wavelengths. The important property of both measures is that they stay constant regardless of the distance from a light source to a sensor (assuming no influence of the medium, in which the light travels). The sensor can be camera's CCD chip or a photoreceptor in the eye. Most HDR images do not contain actual values of luminance, but the values of **luminance factor**. Luminance factor is a relative value of luminance that differs from actual luminance by a constant multiplier.

Although radiometric or photometric units give probably the most accurate description of light, the output of most imaging systems, including photoreceptors, is neither luminance nor spectral radiance. Figure 2 illustrates the response functions that describe the relation between input luminance and output values for several digital cameras. The response of most imaging systems usually follows an S-shaped curve, which tends to saturate both the highest and the lowest luminance values. In this paper we use **luma** term to name the sensor's output values. Luma is a word coined by the NTSC to prevent confusion between the video signal and the traditional meaning of luminance. Since each sensor has its own response characteristic, it is impossible to define a single formula for luma. The relations between luminance and luma used in LDR video compression, which are sometimes called transfer functions, usually involve a power function similar to the gamma correction. These functions however, poorly match performance of the visual system for the luminance values larger than $1\ 000\ cd/m^2$. To address this problem, more robust luma encodings that are suitable for the full range of visible luminance have been proposed.⁵ The encoding that defines luma in terms of sensitivity of the human visual system to light, is shown in Figure 2-right. One advantage of such perceptual representation of luminance is that such a measure of light is perceptually linearized, which means that luma values correlate with our perception of brightness.

Another important aspect for most imaging applications is the range of luminance and luma values that an imaging pipeline should handle. A sufficient range of luminance is within $10^{-5}\ cd/m^2$ and $10^{10}\ cd/m^2$, which can capture the luminance of both a moonless sky ($3 \cdot 10^{-5}\ cd/m^2$) and the surface of the sun ($2 \cdot 10^9\ cd/m^2$). This range of luminance can be encoded as luma using 4096 discrete steps so that the difference between two consecutive steps is not perceivable.⁵ This shows that even if the absolute range of luminance the eye can see is impressive, the actual limitation of the HVS does not allow us to see more than about 4000 visually different shades of gray, and those can be seen only if the slow mechanisms of visual adaptation are involved.

In this short overview we intentionally skipped most aspects of image appearance, such as perception of brightness, hue and colorfulness, as well as the concepts of white point and chromatic adaptation. An excellent discussion of these can be found in.³

1.	Contrast ratio	$1 : R = 1 : \frac{Y_{peak}}{Y_{noise}}$	general, display specifications
2.	log-10 units (orders of magnitude)	$M = \log_{10} \frac{Y_{peak}}{Y_{noise}}$	general
3.	Exposure latitude (f-stops)	$EL = \log_2 \frac{Y_{peak}}{Y_{noise}}$	photography
4.	Signal to noise ratio	$SNR = 20 \cdot \log_{10} \frac{N_{max}}{RMS_{noise}} [db]$	camera specifications
5.	Density range	$DR = D_{max} - D_{min} \approx M$	photography, film scanners

Table 1. Measures of dynamic range used in different disciplines. Y_{peak} is the representative peak (maximum) luminance value and Y_{noise} is the level of noise.

3.2. Dynamic Range

Even more confusing quantity used in digital imaging is the dynamic range. The dynamic range is usually understood as a ratio of the highest and the lowest luminance in an image. However, in most imaging systems the lowest luminance is limited by the noise of that system, such as flare in camera lens, ambient light reflected from a display, or noise in a digital photograph. Therefore, the dynamic range is in practice defined as a ratio of the representative peak signal to the level of noise in an image. For example, if we assume that a computer monitor is almost perfectly black when the pixels are set to zero, which means that luminance of the screen surface is very close to 0 cd/m^2 , the dynamic range of such a theoretical monitor is infinitively high (since the peak luminance is divided by a very small number). However, in real-world the minimum luminance of a good quality LCD monitor in a normally lit room is about 1 cd/m^2 . If the maximum luminance of a LCD computer display is about 300 cd/m^2 , its dynamic range is in fact 1:300.

Camera manufactures usually report the dynamic range using the **signal-to-noise** ratio of an image sensor. Such ratio is measured in decibels using formula 4 in Table 1, where N_{max} is the maximum capacity of a well (given in the number of electrons) and RMS_{noise} is the root mean square of noise. Such measure is usually only a theoretical maximum dynamic range of a camera, which in practice is limited by other camera's elements, such as lens, an A/D converter, and processing performed before an image is stored.

A different measure of dynamic range is used in the photography. The amount of light that passes through lens and reaches a camera's film or digital sensor is expressed as the **f-number** and written as $f/\#$, where $\#$ is the ratio of the focal length and the diameter of the entrance pupil. The sequence of such f-numbers that results in halving the amount of light (luminance) reaching the sensor is a sequence of **f-stops**. The f-stops form a geometric series of powers of $\sqrt{2}$: $f/0.7$, $f/1$, $f/1.4$, $f/2$, $f/2.8$, $f/4$, $f/5.6$, $f/8$, and so on. Therefore, photographers say that a scene has eight f-stops instead of saying that a scene has a dynamic range or contrast ratio 1:256. The number of f-stops is called **exposure latitude** and therefore a high dynamic range image is better known in photography as an image of large exposure latitude (refer to item 3 in Table 1). The best film stocks offer about 12 f-stops of exposure latitude, which corresponds to about 3.5 log-10 units. This is still lower dynamic range than the one that can be captured with HDR cameras or multi-exposure techniques, but it shows that high dynamic range images are not so new to the photography.⁶ Yet another measure of dynamic range that can be found in photography is based on the system of **print zones** introduced by Ansel Adams.⁷ Print zones correspond roughly to f-stop units (they double or halve the amount of captured light), but they are additionally associated with the shades of gray in the resulting print.

The dynamic range measured for analog films is usually expressed as a **density range**. This measure is a difference between the maximum (D-Max) and the minimum (D-Min) tonal values that a film can register (see item 5 in Table 1). Since D-Min and D-Max values are measured on a base-10 log scale, the density range is equivalent to "orders of magnitude" or log-10 units (see item 2 in Table 1). The density range of a good quality film is about 3.4D (note the "D" letter indicating density measure).

All measures of dynamic range discussed in this section and summarized in Table 1. The last remaining aspect is the dynamic range that can be perceived by the human eye. The light scattering on the optic of the eye can effectively reduce the maximum luminance contrast that can be projected onto to retina to 2–3 log-10 units. However, since the eye is in fact highly active sensor, which can rapidly change the gaze and locally adapt, people are believed to be able to perceive simultaneously the scenes of 4 or even more log-10 units [4, Section 6.2]. The precise value of the maximum perceived dynamic range is difficult to measure, as it depends on a scene and the criteria used to assess the visual performance.

4. HIGH DYNAMIC RANGE IMAGING

The main benefit of high dynamic range images is not only high accuracy that matches or exceeds the processing of the human eye. Equally important is the concept of a scene-referred representation of data, which, unlike commonly used output-referred representations, contains enough information to achieve the desired appearance of the scene on a variety of output devices. The scene does not need to be an exact photograph that matches the real-world with photometric accuracy. It could be as well computer graphics rendering (realistic or stylized), or a result of image editing and compositing. Taking a scene-referred image as input, a display device should render it optimally for its capabilities (maximum brightness, dynamic range, color gamut) and viewing conditions (ambient illumination, display size).

High dynamic range imaging promises truly device independent representation of visual content. The need for such a device independent format has become even more important recently, as exchange of digital images, either by Internet or by mobile phones, has become ubiquitous. Also, the output medium is no longer limited to a paper print and a CRT display, but includes LCD, Plasma, LED-backlight, e-paper and OLED displays. The color gamut, dynamic range and maximum brightness can significantly differ between high-end (e.g. LCD TV) and low-end (e.g. mobile phones) displays. Preserving appearance across the variety of displays is almost impossible in such situation. But still, all these displays can show an acceptable image that would preserve the most important image features, if there are provided with sufficient information on an the image content, encoded, for instance, in an HDR image. Moreover, the technological limitations, which were major motivation for output-referred formats (e.g. gamma correction matching the characteristic of most TV sets), no longer exist and potential bit savings coming from images profiled for displays are much less important. Display devices are no longer intended to display the content as it is, but they also apply complex image processing to enhance the resulting image. However, to fully utilize capabilities of the display and its display algorithms, the input content must be device-independent and of possible high quality, exceeding the standards of traditional imaging. High dynamic range image representation meets both of these requirements and therefore offers a good alternative to the traditional image representation.

5. SOFTWARE FOR PROCESSING HDR CONTENT

Although there is a plentitude of software that can process traditional images and video, very few programs support high dynamic range content. Even if imaging software offers higher bit-depth precision, usually up to 16 bits per color channel, it still relies on the assumption that the brightest possible color is white and the dynamic range of a scene is limited. HDR images and video are most naturally represented as floating point numbers, and they often contain bright highlights or light sources that exceed the luminance of white diffuse surfaces and thus the largest code value used in traditional imaging.

One of the first software fully supporting HDR content was Radiance ray-tracing package. Since the ray-tracing methods, which simulate the distribution of light, are slow to compute and can give accurate values of spectral radiance, it would be waste to store only partial information in a traditional low-dynamic range image. To accurately preserve the result of many hours of computations, one of the first HDR formats was developed, known as Radiance RGBE format⁸ and recognized by the extension *.hdr* or *.pic*. This format, as well as other HDR formats such as logLuv TIFF, had been used almost exclusively for storing results of computer graphics rendering algorithms, until HDR photography was popularized by multi-exposure methods. In 2003 Industrial Light and Magic released and started supporting OpenEXR library for storing images in an HDR format,⁹ recognized by the extension *.exr*. The library is intended to establish an industry standard and a unified format for storing HDR content. The library offers many advanced features, such as arbitrary number of channels (for storing depth, motion flow or object IDs), user tags, an efficient 16-bit floating point format and lossless wavelet compression. Only recently, major software vendors have released or announced the release of their products supporting HDR content, including the most popular Radiance RGBE and OpenEXR formats.

Driven by the scarcity of HDR-enabled software, we have created the *pfstools* package*, which is a collection of several command line tools for reading, writing, processing and viewing HDR images and video. The software was intended to support our current research projects, but we found it reliable and useful enough to release it as open source project under the GPL license. Since the release, we have found that *pfstools* have been used for high definition HDR video encoding, medical imaging applications, variety of tone mapping projects, texture manipulations and quality evaluation of computer graphics rendering.

*Home page of *pfstools* can be found at <http://pfstools.sourceforge.net/>.

To assure maximum flexibility, we have designed *pfstools* based on the successful solutions found in several other image processing software packages. All programs that are part of *pfstools* exchange HDR images or frames using a generic and simple image format, which requires only a few lines of code to read or write. The format is based on the following assumptions:

- Pixel values are stored with high precision as 32-bit floating point numbers. Such precision exceeds capabilities of the human visual system⁵ and is sufficient for most applications. 32-bit floating point format is natively supported on most of platforms;
- Images can contain an arbitrary number of channels (layers), which can represent not only color, but also depth, alpha, and texture attributes;
- User data entries can store additional, application specific information, such as colorimetric coordinates of the white point or calibration information. These are stored using *user tags*, which are name and value pairs;
- There is no compression as the files in this format are intended to be transferred internally between applications without writing them to a disk.
- A sequence of images is interpreted by all “pfs-compliant” applications as animation frames, so that video can be processed in the same way as images.

A few channels in the *pfs* format have a predefined function. For example, channels with the IDs X, Y and Z are used to store color data in the CIE XYZ (absolute) color space. This is different to most imaging frameworks that operate on RGB channels. The advantage of the CIE XYZ color space is that it is precisely defined in terms of spectral radiance while any RGB color space must specify at least its primaries (usually given as XYZ coordinates) to allow conversion between color spaces. Additionally, when the primaries are badly chosen, the RGB color coordinates can be negative, which can not be handled properly by a number of image processing algorithms. The format is described in detail in a separate specification[†].

pfstools are a set of command line tools with almost no graphical user interface. This greatly facilitates scripting and lessens the amount of work needed to program and maintain a user interface. The exception is a viewer of HDR images. The main components of *pfstools* are: programs for reading and writing images in all major HDR and LDR formats (e.g. OpenEXR, Radiance’s RGBE, logLuv TIFF, 16-bit TIFF, PFM, JPEG, PNG, etc.), programs for basic image manipulation (rotation, scaling, cropping, etc.), an HDR image viewer, and libraries that simplify file format reading and writing for C++, Java, matlab and GNU Octave (high level mathematical language similar to matlab). The *pfstools* framework does not impose any restrictions on the programming language. All programs that exchange data with *pfstools* must read or write the file format, but there is no need to use any particular library. Additional functionality is provided by separate packages, which are described in the next sections.

The typical usage of *pfstools* involves executing several programs joined by UNIX pipes. The first program transmits the current frame or image to the next one in the chain. The final program should either display an image or write it to a disk. Some examples of command lines are given below:

```
pfsin_input.exr|_pfsfilter|_pfsout_output.exr
```

Read the image `input.exr`, apply the filter `pfsfilter` and write the output to `output.exr`.

```
pfsin_input.exr|_pfsfilter|_pfsview
```

Read the image `input.exr`, apply the filter `pfsfilter` and show the result in an HDR image viewer.

```
pfsin_in%04d.exr--frames_100:2:200_\n|_pfsfilter|_pfsout_out%04d.hdr
```

Read the sequence of OpenEXR frames `in0100.exr`, `in0102.exr`, ..., `in0200.exr`, apply the filter `pfsfilter` and write the result in Radiance’s RGBE format to `out0000.hdr`, `out0001.hdr`, ...

[†]Specification of the pfs format can be found at: http://www.mpi-sb.mpg.de/resources/pfstools/pfs_format_spec.pdf

6. ACQUISITION

In this section we focus on an entry point to the HDR pipeline which is acquisition of HDR images. The goal of HDR acquisition methods is to achieve scene referred HDR representation rather than generally used display (output) referred representation (with the dynamic range limited to the capabilities of low dynamic range monitors).

Images of an extended dynamic range can be captured using standard devices, such as professional digital cameras that can store images in the RAW formats. The dynamic range of the RAW formats is not limited due to image processing in a camera (like sharpening which reduces the dynamic range about an exposure stop) so more accurate information about scene luminance can be achieved (refer to Section 3.2 for details on f-stops definition). Another example are professional film scanners used in cinematography, which can digitalize film density (logarithmic characteristic) with 12-bit precision.

The devices that can capture high dynamic range data are used in various fields like ultrasonography, radiometry, seismology, etc. In typical imaging pipelines the dynamic range is compressed to display referred representation. It causes significant limitations in image analysis, processing and displaying of data. In the future, the above mentioned devices will tend to extend their dynamic range to increase their accuracy.

There are already devices on the market that are natively designed to capture HDR. The HDR cameras, whose survey can be found in [4, Section 4.9], exploit two major methods to extend the dynamic range of a sensor and retain details in dark and bright regions of a scene. The first method employs the logarithmic response of a CMOS sensor to compute the logarithm of an irradiance in the analog domain.¹⁰ The second method involves measuring the time that a sensor pixel takes to attain full capacity,¹¹ which is proportional to spectral radiance. This way a sensor never saturates and higher dynamic range can be captured.

HDR images and video can be also generated using computer graphics techniques. Both real time (GPU based) and off-line rendering architectures support floating point color representation.

A number of high dynamic range acquisition techniques involves capturing multiple images of the same scene using different exposures and then merging these images into an HDR image. Beam splitters can be used to expose a few sensors of a different sensitivity simultaneously. The patterns that change the exposure of individual pixels of a sensor enable to extend dynamic range by decreasing image resolution.¹² Similar solution was proposed in^{13,14} where the exposure of individual sensor pixel can adapt to scene radiance by using an optical attenuator. The transmittance of the attenuator is computed based on the brightness measured by the pixel. These methods, however, require specialized hardware. In the following section, we discuss the technique for capturing HDR images using typical low dynamic range cameras.

6.1. Image acquisition from multiple exposures

High dynamic range images can be captured using a sequence of photographs of the same scene taken with different exposure using so called multi-exposure technique. This technique allows to increase the dynamic range and to reduce the noise in an image. The values of luminance recovered using multi-exposure techniques are not meant to be exact measurements, since they can be severely distorted by camera optics.¹⁵ The accuracy of reconstructed luminance values depends on a scene, with better accuracy for the scenes containing fewer number of bright pixels.

The goal of the method is to vary the exposure in order to control the light levels to be captured. By increasing the exposure time, low-light details will be represented well, at the cost of losing information in highly illuminated areas. On the other hand, decreasing exposure time will show the details in areas of high illumination (see Figure 3). The sequence of photographs is then merged to create an HDR image containing the values approximately proportional to the luminance of a real scene. Such luminance factors (refer to Section 3.1) need to be multiplied by a constant value, which depends on a camera and lens, to get actual level of luminance. Such constant number can be easily found if we can measure the luminance of a photographed surface.¹⁶ Under the assumption that the camera is perfectly linear, each photograph can be brought into the same domain by dividing each pixel by the exposure value and then merged by weighted addition. Unfortunately, typical cameras are not linear in the most cases and a *camera's response function* should be measured in order to linearize the data.

A *camera response function* defines the relation between pixels values and the intensity of light that reaches camera's CCD sensor. There are several methods of estimating camera response function.¹⁷⁻²² Figure 2(left) shows an example response function of several digital cameras. The procedure of taking the input photographs used for capturing the characteristics is presented in [4, Section 4.6].



Figure 3. Four LDR images selected from a sequence of photographs used to build an HDR image. The resulting HDR image (right) was created by the *pfshdrcalibrate* and, to display in LDR media, tone mapped using contrast domain operator.²³

The *pfshdrcalibrate* program in *pfscalibration* package can merge a sequence of standard photographs to create an HDR image. The *pfscalibration* package is integrated with *pfstools* and is available with the source code under the GPL license[‡]. The merging algorithm is based on the technique proposed in.¹⁹ The *pfshdrcalibrate* reads input pictures stored in various formats (e.g. JPEG, TIFF or RAW). The exposure parameters of input images are either extracted from the file header or can be entered manually. The result of combining the photographs into an HDR image are shown in Figure 3. The *pfshdrcalibrate* program can be also used to estimate a camera response function (see¹⁶ for details).

7. IMAGE RENDERING – TONE MAPPING

	tone mapping algorithm	type	video	real-time	key features
1.	Adaptive Logarithmic Mapping ²⁴	global	no*	yes	brightness mapping
2.	Time-Dependent Visual Adaptation ²⁵	global	yes	yes	psychophysical model of visual adaptation
3.	Photoreceptor ²⁶	global	no*	yes	photoreceptor model with chromatic adaptation (von Kries)
4.	Photographic Tone Reproduction ⁶	local	no*	yes	film response, zone system, detail enhancement
5.	Bilateral Filtering for HDR ²⁷	local	no	no	contrast reduction in illumination layer while details preserved
6.	Local adaptation TM ²⁸	local	no	no	modeling HVS response to luminance (based on t.v.i. function)
7.	Gradient Domain Compression ²⁹	local	no	no	local attenuation of large scale gradients
8.	Lightness Perception ³⁰	local	no	no	computational model of lightness perception
9.	Contrast Mapping and Equalization ²³	local	no	no	processing of image contrast
10.	Histogram Equalization ³¹	global	no	no	histogram equalization with perceptual constraints

Table 2. Tone mapping operators implemented in *pfstmo* package and their characteristics. Refer to Section 7 for details.

Standard 8-bit images can be displayed literally because of their display referred representation. The HDR images, however, contain a scene referred representation and need to be processed prior to display. In particular, most of current displays have insufficient capabilities to display the exact luminance values stored in an HDR image because the dynamic range of these devices is limited. In this section we are concerned with the depiction of HDR images on devices with various capabilities and given various application requirements – *tone mapping*.

By definition, tone mapping algorithm is a function that transforms luminance to a displayable range of luma values. In the case when tone mapping takes into account color information, the color channels linearly related to luminance are transformed to red, green and blue channels of pixel intensity values.

Generally, the tone mapping algorithms are derived from either brightness mapping functions (similar to response of an analog film),^{6,24} threshold visibility functions,^{23,28,31} physiological behavior of human eye,²⁵ contrast theories or lightness

[‡]The home page of *pfscalibration* package can be found at: <http://www.mpii.mpg.de/resources/hdr/calibration/pfs.html>

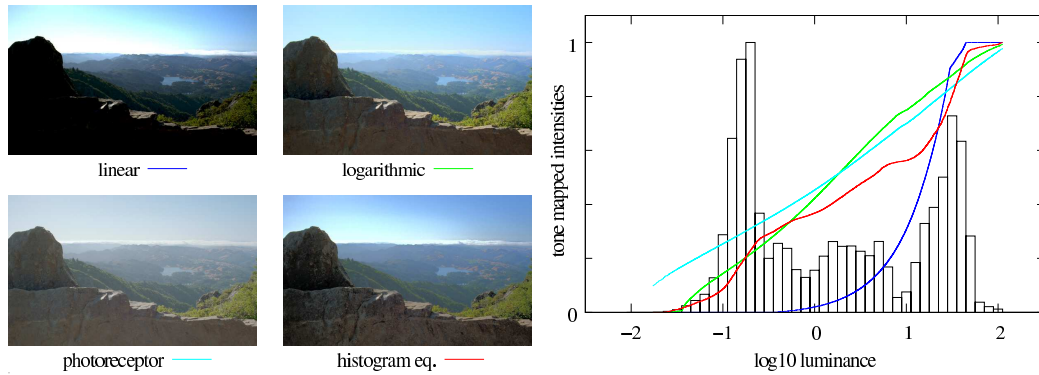


Figure 4. Various mapping curves of global tone mapping operators adjusted for a sample HDR image. The histogram in the background illustrates the luminance of the sample HDR image.

perception theories.³⁰ The development of each of these methods was driven by requirements of a particular application and so far a universal tone mapping algorithm is not available. Tone mapping can be profiled to achieve different goals: an automatic method producing *nice looking* images, an emphasis on details enhancement for scientific visualizations, a perceptual match between a real world scene and a displayed image, depiction of stylized impression of a scene.³² Despite the vast diversity of natural and synthetic HDR images and the possible luminance values inaccuracy found in HDR photographs, a robust algorithm is expected to provide consistent results while avoiding undesirable artifacts including perceivable contrast reversals and halos along high contrast edges. The algorithm is in most cases expected to be automatic with few intuitive parameters that provide possibility for fine tuning. It should address different capabilities of display devices and potential differences in observation conditions. Additionally, the tone mapping of video sequences must be coherent over time. Apparently not many algorithms are time coherent, however most methods can achieve time coherence by incorporating a model of visual adaptation.²⁵

Tone mapping can in principle produce a very good depiction of an HDR image even though the display is incapable to reproduce appropriate range of luminance, because the human visual system is generally not able to estimate absolute luminance and is sensitive to spatial patterns in the luminance differences instead.³³ A thoughtful manipulation of relative luminance to create desired global and local contrasts can compensate for limited dynamic range, difference in adaptation level of observers and its influence on perceivable contrast, ambient illumination, and particular display characteristics: LCD, CRT, projector, plasma, HDR.²

Certain visual phenomena are however related to absolute levels of luminance and when the luminance is treated in a relative way (see the discussion of *luminance factor* in Section 3.1) these effects may disappear. Humans are particularly aware of loss of color visibility and visual acuity for low luminance levels, and the glare effect for strong light sources. This phenomena are well measured in psychophysics and appropriate models are available, therefore some tone mapping operators simulate such effects after performing the tone mapping^{31,34} to obtain a better perceptual match to the original scene. Such effects can be also applied in real time.³⁵

A selection of the state-of-the-art tone mapping operators have been implemented within *pfstools* as the *pfstmo* package. They are summarized in Table 2. In Figure 4 we show for comparison several tone mapping curves adjusted to a sample HDR image. Clearly, linear scaling of luminance cannot map the input dynamic range to displayable values and the details in high and low luminance ranges are lost. The full range of input luminance values can be successfully mapped using the adaptive logarithmic mapping.²⁴ Equally good performance can be achieved with the photoreceptor model,²⁶ though the image will be considerable brighter as the mapping curve maps low luminance values to higher display values. The curve of the histogram adjustment method³¹ not only covers the whole input range but also fine tunes its shape to assign wider displayable range to highly populated luminance ranges that leads to a well optimized final result. Further improvements can be achieved with *local* algorithms which can emphasize pixel visibility by considering the neighboring pixels.^{6,23,28,30}

The scene referred representation of image and video data has an obvious advantage given that the depiction of HDR content can be highly customized just before the display. The visual quality of displayed content can be seriously degraded when already tone mapped images and video are stored and transmitted without any prior knowledge of their actual

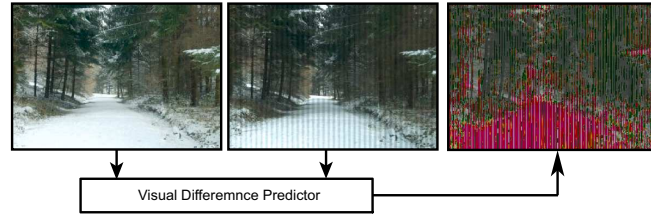


Figure 5. High Dynamic Range Visual Difference Predictor (HDR-VDP) takes as an input two images and produces as a result the probability of detection map. Such a map tells how likely a human observer will notice a difference between these two images in particular part of a scene.

display conditions or target application. Furthermore, perceptual effects which are usually not evoked due to the change in absolute luminance can be added as post processing. Unfortunately, tone mapping algorithms are generally difficult to evaluate. Such evaluation is possible only by conducting tedious subjective comparison experiments and the results of such experiments are often inconclusive.^{36,37}

pfstmo package is integrated with *pfstools* and is available with the source code under the GPL license[§]. The data stream send to a tone mapping algorithm is scene referred in the HDR pipeline, but the output stream is already display referred. In practice this means that tone mapping is the end point of an HDR pipeline.

8. FIDELITY AND QUALITY ASSESSMENT

Visual difference metrics can predict whether differences between two images are visible to the human observer or not. Such metrics are used for testing either visibility of information (whether we can see important visual information) or visibility of noise (to make sure we do not see any distortions in images, e.g. due to inaccurate CG rendering or lossy compression).

Figure 5 illustrates how two input images, a reference image (left) and a distorted image (middle), are processed with the visual difference predictor to produce a probability of detection map (right). Such probability of detection map tells how likely we will notice a difference between two images for each part of a scene. Red color denotes high probability, green - low probability. Red color is mostly present in the areas where there is a snow covered path. Because of smooth texture of the snow, there is not much visual masking and distortions are easily visible.

High Dynamic Range Visible Difference Predictor (HDR-VDP)³⁸ is the implementation of a difference predictor that can work within the complete range of luminance the human eye can see. An input to our metric is an HDR image, or an ordinary 8-bits-per-color image, converted to the actual luminance values. The images must be properly calibrated in absolute luminance units in order to guarantee reliable prediction. This is because the metric takes into account the aspects of high contrast vision, like scattering of the light in the optics (OTF), nonlinear response to light for the full range of luminance, and local adaptation. These have varied influence on the detection threshold depending on the illumination level of the scene. For example, small details are not be noticeable under very dim light (scotopic vision) but are well visible when they are lit by sunlight (photopic vision).

HDR-VDP is available for the research community with the source code under the GPL license[¶]. It is integrated with *pfstools*, so that it can take as an input any pair of images that can be read by this package.

9. CONCLUSIONS

The advances in display and sensor technology have revealed the limitation of the existing low-dynamic range imaging pipelines. Those limitations can be addressed and eliminated in the HDR imaging pipelines that offer higher precision of the visual data. In this work we outline the fundamental quantities associated with HDR images. We introduce a concept of an universal imaging framework in which images are represented in a generic format that fulfills the requirement of variety

[§]The home page of *pfstmo* package can be found at: <http://www.mpii.mpg.de/resources/tmo/>

[¶]HDR-VDP home page can be found at: <http://hdrvdp.sourceforge.net/>

of applications. Such requirements include photometric accuracy of data and flexibility achieved by an arbitrary number of layers and user data entries. Finally, we describe the software packages for acquisition, tone mapping and comparing HDR images that we have made available for research community. We believe that the software can be a valuable help for all those who are involved in HDR imaging.

REFERENCES

1. C. Bartleson, "Memory colors of familiar objects," *Journal of the Optical Society of America* **50**(1), pp. 73–77, 1960.
2. H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," *ACM Trans. on Graph.* **23**(3), pp. 757–765, 2004.
3. R. Hunt, *The Reproduction of Colour in Photography, Printing and Television: 5th Edition*, Fountain Press, 1995.
4. E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging. Data Acquisition, Manipulation, and Display*, Morgan Kaufmann, 2005.
5. R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Lossy compression of high dynamic range images and video," in *Human Vision and Electronic Imaging XI, SPIE, volume 6057*, 2006.
6. E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. on Graph.* **21**(3), pp. 267–276, 2002.
7. A. Adams, *The Print, The Ansel Adams Photography Series 3*, New York Graphic Society, 1981.
8. G. Ward, "Real pixels," *Graphics Gems II*, pp. 80–83, 1991.
9. R. Bogart, F. Kainz, and D. Hess, "OpenEXR image file format," in *ACM SIGGRAPH 2003, Sketches & Applications*, 2003.
10. U. Seger, H.-G. Graf, and M. Landgraf, "Vision assistance in scenes with extreme contrast," *IEEE Micro* **12**(1), pp. 50–56, 1993.
11. V. Brajovic and T. Kanade, "A sorting image sensor: An example of massively parallel intensity-to-time processing for low-latency computational sensors," in *Proceedings of ICRA-96, Minneapolis, MN, 22-28 Apr*, pp. 1638–43, 1996.
12. S. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 472–479, 2000.
13. S. Nayar and V. Branzoi, "Adaptive dynamic range imaging: Optical control of pixel exposures over space and time," in *Proc. of IEEE International Conference on Computer Vision (ICCV 2003)*, pp. 1168–1175, 2003.
14. S. Nayar, V. Branzoi, and T. Boult, "Programmable imaging using a digital micromirror array," in *CVPR04*, pp. I: 436–443, 2004.
15. J. McCann and A. Rizzi, "Spatial comparisons: the antidote to veiling-glare limitations in HDR images," in *Third Americas Display Engineering and Applications Conference (ADEAC)*, 2006.
16. G. Krawczyk, M. Goesele, and H.-P. Seidel, "Photometric calibration of high dynamic range cameras," Research Report MPI-I-2005-4-005, Max-Planck-Institut für Informatik, April 2005.
17. P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of ACM SIGGRAPH 1997*, pp. 369–378, 1997.
18. T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," pp. 374–380, 1999.
19. M. Robertson, S. Borman, and R. Stevenson, "Dynamic range improvement through multiple exposures," in *Proceedings of the 1999 International Conference on Image Processing (ICIP-99)*, pp. 159–163, (Los Alamitos, CA), Oct. 24–28 1999.
20. M. Grossberg and S. Nayar, "Determining the camera response from images: What is knowable?," *PAMI* **25**, pp. 1455–1467, November 2003.
21. Y. Tsin, V. Ramesh, and T. Kanade, "Statistical calibration of the CCD imaging process," *Proc. of International Conference on Computer Vision (ICCV)*, pp. I: 480–487, 2001.
22. S. Lin, J. Gu, S. Yamazaki, and H. Shum, "Radiometric calibration from a single image," in *CVPR04*, pp. II: 938–945, 2004.
23. R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," in *APGV '05: 2nd Symposium on Applied Perception in Graphics and Visualization*, pp. 87–94, 2005.
24. F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum, proceedings of Eurographics 2003* **22**(3), pp. 419–426, 2003.

25. S. Pattanaik, J. Tumblin, H. Yee, and D. Greenberg, "Time-dependent visual adaptation for realistic image display," in *Proceedings of ACM SIGGRAPH 2000*, pp. 47–54, 2000.
26. E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Transactions on Visualization and Computer Graphics* **11**(1), pp. 13–24, 2005.
27. F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. on Graph.* **21**(3), pp. 257–266, 2002.
28. M. Ashikhmin, "A tone mapping algorithm for high contrast images," in *Rendering Techniques 2002: 13th Eurographics Workshop on Rendering*, pp. 145–156, 2002.
29. R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. on Graph.* **21**(3), pp. 249–256, 2002.
30. G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Lightness perception in tone reproduction for high dynamic range images," *Computer Graphics Forum* **24**(3), 2005.
31. G. Ward 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**(4), pp. 291–306, 1997.
32. K. Smith, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Beyond tone mapping: Enhanced depiction of tone mapped hdr images," *Computer Graphics Forum* **25**(3), 2006.
33. B. Wandell, *Foundations of Vision*, Sinauer Associates, Inc., 1995.
34. J. Ferwerda, S. Pattanaik, P. Shirley, and D. Greenberg, "A model of visual adaptation for realistic image synthesis," in *Proceedings of SIGGRAPH 96, Computer Graphics Proceedings, Annual Conference Series*, pp. 249–258, Aug. 1996.
35. G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Perceptual effects in real-time tone mapping," in *SCCG '05: Proceedings of the 21st spring conference on Computer graphics*, pp. 195–202, ACM Press, (New York, NY, USA), 2005.
36. F. Drago, W. Martens, K. Myszkowski, and H.-P. Seidel, "Perceptual Evaluation of Tone Mapping Operators with Regard to Similarity and Preference," Technical Report MPI-I-2002-4-002, Max-Planck-Institut fuer Informatik, Oct. 2002.
37. P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," *ACM Transactions on Graphics* **24**(3), pp. 640–648, 2005.
38. R. Mantiuk, S. Daly, K. Myszkowski, and H.-P. Seidel, "Predicting visible differences in high dynamic range images - model and its calibration," in *Human Vision and Electronic Imaging X, SPIE, volume 5666*, pp. 204–214, 2005.