A review of tone reproduction techniques

Kate Devlin
Department of Computer Science
University of Bristol

November 2002

Abstract

The ultimate aim of realistic graphics is the creation of images that provoke the same responses that a viewer would have to a real scene. While research into ways of rendering images provides us with better and faster methods, we do not necessarily see their full effect due to limitations of the display hardware. The low dynamic range of a standard computer monitor requires some form of mapping to produce images that are perceptually accurate. Tone reproduction operators attempt to replicate the effect of real-world luminance intensities. This paper reviews the work to date on tone reproduction techniques. It includes an investigation into the need for accurate tone reproduction and a discussion of techniques to date. The future of tone reproduction is considered, together with the implications of advances in display hardware.

1 Introduction

The ultimate aim of realistic graphics is the creation of images that provoke the same response and sensation as a viewer would have to a real scene, i.e. the images are physically or perceptually accurate when compared to reality. However, realistic rendering is not enough to ensure perceptual fidelity. Displaying an image is also an important part of the overall process, and weaknesses in this area may significantly detract from advances made in image creation.

This paper reviews the state of the art in tone reproduction techniques. Section 2 describes the need for, and the theory behind, tone reproduction. Previous work on tone mapping is discussed in Section 3. The final section, Section 4, examines the potential for future work in this field and the implication of advances in display technology.

2 Tone mapping: an overview

While research into ways of creating images provides us with better and faster methods, we may not see the full effect of these techniques due to display limitations. For accurate image analysis and comparison with reality, the display image must bear as close a resemblance to the original image as possible. Ideally, if a scene in the real world and an image representing that scene (be it computer generated or photographed) are viewed under the same conditions, it is expected that the real-world scene and the image should have the same tones, i.e. the luminance levels of both scenes match.

Physical accuracy alone of an image does not ensure that the scene in question will have a realistic visual appearance when it is displayed. This is due to the shortcomings of standard display devices, which can only reproduce a range of luminance of about 100:1 candelas per square metre (cd/m^2), as opposed to human vision which ranges from 100 000 000:1, from bright sunlight down to starlight (Figure 1). The human eye can accommodate a luminance range of approximately 10 000:1 in a single view, and an observer's adaptation to their surroundings (where their response to a scene changes over time) also needs to be taken into account.

The ratio between the maximum and the minimum tonal values in an image is known as the *dynamic range*. It is this high dynamic range (HDR) that exists in the real world that needs to be scaled in some way to fit a display device that is only capable of outputting a low dynamic range.

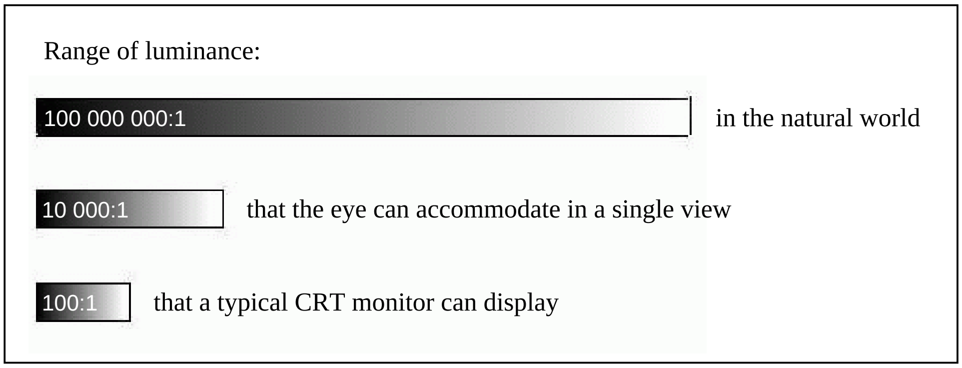

Figure 1: A comparative view of dynamic range.

Tone reproduction (also known as tone mapping) provides a method of scaling (or mapping) luminance values in the real world to a displayable range. Tone reproduction is necessary to ensure that the wide range of light in a real-world scene is conveyed on a display with limited capabilities. In addition to compressing the range of luminance, it can be used to mimic perceptual qualities, resulting in an image which provokes the same responses as someone would have when viewing the scene in the real world. For example, a tone reproduction operator may try to preserve aspects of an image such as contrast, brightness or fine detail — aspects that might be lost through compression.

2.1 The need for accurate tone reproduction

The way in which we perceive images depends on the amount of light available. In dark scenes our visual acuity — the ability to resolve spatial detail — is low and colours cannot be distinguished. This is due to the two different types of photoreceptor in the eye: rods and cones. It is the rods that provide us with achromatic vision at these *scotopic* levels, functioning within a range of 10^{-6} to 10 cd/m^2. Visual adaptation from light to dark is known as *dark adaptation*, and can last for tens of minutes; for example, the length of time it takes the eye to adapt at night when the light is switched off. Conversely, *light adaptation*, from dark to light, can take only seconds, such as leaving a dimly lit room and stepping into bright sunlight. The cones are active at these *photopic* levels of illumination, covering a range of 0.01 to $10^{8} cd/m^2$. The overlap (the *mesopic* levels), when both rods and cones are functioning, lies between 0.01 to 10 cd/m^2. The range normally used by the majority of electronic display devices (cathode ray tubes, or CRTs) spans from 1 to 100 cd/m^2 . More detailed information on visual responses with regard to tone reproduction can be found in the papers by Ferwerda et al. , Pattanaik et al. , and Tumblin [10, 27, 28, 39].

Despite a wealth of psychophysical research, our knowledge of the Human Visual System (HVS) is still limited, but its ability to perceive such a wide dynamic range in the real-world requires some method of reproduction that produces similar images on display devices. In situations where predictive imaging is required, tone reproduction is of great importance to ensure that the conclusions drawn from a simulation are correct (Figure 2).

Changes in the perception of colour and of apparent contrast also come into play when mapping values to a display device. The development of new psychophysically-based visual models seeks to address these factors. Often methods of tone mapping tend to concentrate on singular aspects for singular purposes, such as brightness matching, or visibility preservation. This approach is understandable given the deficit in HVS knowledge, but is inefficient as the HVS responds as a whole, rather than as isolated functions. New psychophysical research is needed to address the workings of the HVS in their totality. It is this demand to present the aspects of a real-world scene on a limited medium that makes tone reproduction such a challenging task.

2.2 Tone mapping: art, television and photography

Tone mapping was developed for use in television and photography, but its origins lie in the field of art where artists make use of a limited palette to depict high contrast scenes. It takes advantage of the fact that the HVS has a greater sensitivity to relative rather than absolute luminance levels [14].

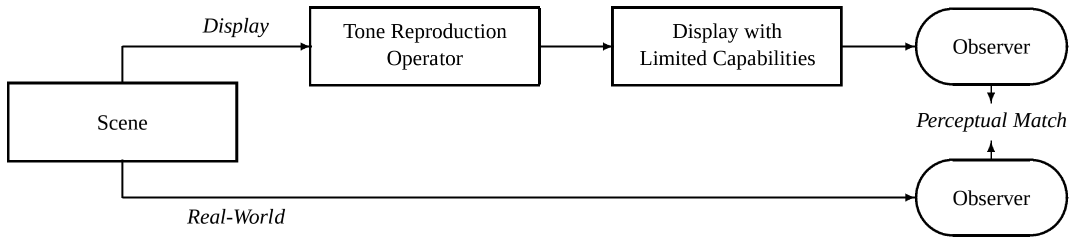

Figure 2: Ideal tone reproduction process

Tone reproduction is already used extensively to good effect in photography and television, and some of the methods used computer graphics have been inspired and influenced by techniques in these media. The use of tone reproduction in photography and television today is explained in Hunt's "The Reproduction of Colour in Photography, Print and Film" [14] and Poynton's "A Technical Introduction to Digital Video" [30], which give comprehensive explanations on the subject. This research area is outside the scope of this paper, and it is suggested that readers with an interest in this area refer initially to these works.

2.3 Gamma correction

The relationship between the input voltages of a CRT device and the output intensity of the display is non-linear. For CRTs, the use of RGB values to express colour is actually specifying the voltage that will be applied to each electron gun. The luminance generated is not linearly related to this voltage. The intensity displayed on a CRT is proportional to the input voltage raised to the power gamma (γ), so that if a pixel is to have an intensity of x, the displayed intensity will actually be x^{γ}. For most CRTs gamma is a power of 2.5, although the actual value of the exponent varies [29].

Gamma correction is a way of compensating for these discrepancies, and is implemented by applying the inverse of the voltage-to-intensity signal, i.e. $x = x^{\frac{1}{\gamma}}$, to the RGB data before display.

Gamma correction should always be considered as a step towards displaying an image as it was intended to be seen. However, although gamma correction goes some way towards correcting the data, there is still scope for variation. Most monitors provide brightness and contrast controls. Correction may also have been applied to the image data or in the user software. These potential areas for correction can lead to inconsistencies and it cannot be assumed that an approximation of an ideal display has been achieved.

3 Previous work on Tone Reproduction

Reviews of tone reproduction operators have been carried out in previous years [19, 21], and these also examine the HVS factors that influence the techniques.

Two types of tone reproduction operators can be used: *spatially uniform* (also known as *single-scale* or *global*) and *spatially varying* (also known as *multi-scale* or *local*). Spatially uniform operators apply the same transformation to every pixel. A spatially uniform operator may depend upon the contents of the image as a whole, as long as the same transformation is applied to every pixel. Conversely, spatially varying operators apply a different scale to different parts of an image. A further aspect to tone reproduction is time. It should be noted that the above definitions do not account for temporal differences (such as adaptation over time), so we have included these under a separate category of *time dependent* tone reproduction operators.

This section aims to provide an overview of the tone reproduction methods that have been published to date. Figure 3 shows the categorisation and development of tone reproduction methods and Table 1 gives an overview of the attributes of tone reproduction methods.

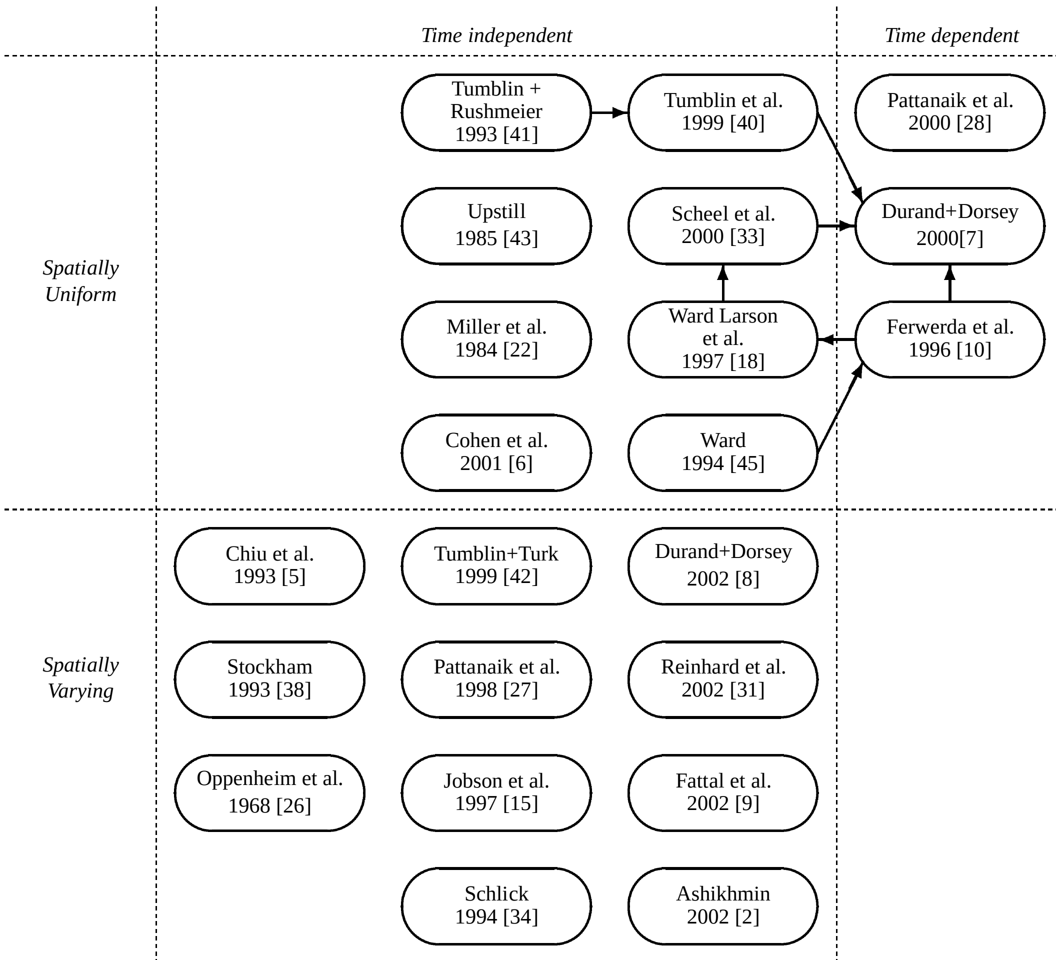

Figure 3: Categorisation and development of Tone Reproduction Methods

3.1 Spatially uniform operators

In 1984 **Miller, Ngai and Miller** [22] were the first to use experimental data to try to match brightness in a real scene to brightness of a displayed image of that scene, for the purpose of determining pixel luminance for their architectural rendering system [1]. They used psychophysical data on brightness perception from work by Stevens and Stevens [37]. **Upstill's** 1985 PhD thesis reinforced the need for perceptual tone reproduction through the use of an explicit perceptual model [43].

Tumblin and Rushmeier [41], also focused on preserving the viewer's overall impression of brightness, providing a theoretical basis for perceptual tone reproduction, again by using Stevens and Stevens data.. This model of brightness perception is not valid for complex scenes but was chosen by Tumblin and Rushmeier due to its low computational costs. Their aim was to create a 'hands-off' method of tone reproduction in order to avoid subjective judgements. They created observer models — mathematical models of the HVS that include light-dependent visual effects while converting real-world luminance values to perceived brightness images. The real-world observer corresponds to someone immersed in the environment, and the display observer to someone viewing the display device. Their tone reproduction operator converts the real-world luminances to the display values, which are chosen to match closely the brightness of the real-world image and the display image. If the display luminance falls outside the range of the frame-buffer then the frame-buffer value is clamped to fit this range.

This method is limited to greyscale and by the preservation of brightness at the expense of visibility in high dynamic scenes [18]. It has been noted that as the operator can handle extremes of brightness, some images tend to appear too dark but this may work in its favour if the analysis of extreme lighting conditions is required [19].

Ward's model [45] dealt with the preservation of perceived contrast rather than brightness. Ward aimed to keep computational costs to a minimum by transforming real-world luminance values to display values through a scaling factor, concentrating on small alterations in luminance that are discernible to the eye. Based on a psychophysical contrast sensitivity model by Blackwell [3] he exploited the fact that the consequence of adaptation can be regarded as a shift in the absolute difference in luminance required for the viewer to notice the variation. Blackwell produced a comprehensive model of changes in visual performance due to adaptation level. This means that a Just Noticeable Difference (JND) in the real-world can be mapped as a JND on the display device.

This approach is useful for displaying scenes where visibility analysis is crucial, such as emergency lighting, as it preserves the impression of contrast. It is also less computationally expensive than Tumblin and Rushmeier's operator but the use of a linear scaling factor causes very high and very low values to be clamped and correct visibility is not maintained throughout the image [18]. It should also be noted that Blackwell's experiments were conducted in near-perfect laboratory conditions and therefore do not take into consideration the complexities of typical workplace viewing conditions.

Further work by **Ward Larson, Rushmeier and Piatko** [18] presented a histogram adjustment technique for reproducing perceptually accurate tones in high dynamic display scenes, extending earlier work by Ward [45] and Ferwerda et al. [10]. The main focus of this work was object visibility and image contrast, with a secondary goal of recreating the viewer's subjective response so that their impression of the real and virtual scenes were consistent [18].

This technique employs the knowledge that the eye is sensitive to relative rather than absolute changes to luminance, so bright ares should be displayed as bright and dim areas as dim, irrespective of the actual absolute luminance intensity values. Luminance levels are not constant across an image, but appear in clusters that vary in intensity. Also, the eye adapts rapidly to a 1° visual field around the fixation point. For these reasons, Ward Larson et al. 's operator makes adjustments on the basis of luminance adaptation levels in an image rather than on spatial location.

The field of image processing has developed methods to adjust image contrast and visibility. One such method is the histogram equalisation technique whereby the grey levels in an image are redistributed to make better use of the display device range and maximise visibility and contrast. Ward Larson et al. exploit this idea of altering histograms and using perceptual models to guide alteration, with their aim being to simulate, rather than maximise, visibility in an image.

A log of luminances averaged over 1° areas (which correspond with foveal adaptation levels for possible points in an image) is obtained, and a histogram and cumulative distribution function is built from this information. Cumulative distribution of the luminance histogram is used to identify clusters of luminance levels and initially map them to the display values using a histogram adjustment technique that is based on human contrast sensitivity. Ferwerda et al. 's [10] threshold sensitivity data is used to compress the original dynamic range to that of the display device, subject to the contrast sensitivity limitations of the eye. Although this method is described here as spatially uniform, spatial variation is introduced through the use of models for glare, acuity and chromatic sensitivity to increase perceptual fidelity.

In 1999 **Tumblin, Hodgkins and Guenter** [40] produced two new tone reproduction operators by imitating some of the HVS's visual adaptation processes, and also revised Tumblin and Rushmeier's [41] earlier work. The first, a layering method, builds a display image from several layers of lighting and surface properties. This is done by dividing the scene into layers and compressing only the lighting layers while preserving the scene reflectances and transparencies, thus reducing contrast while preserving image detail. Their compression function follows the work of Schlick [34]. This method only works for synthetic images where layering information from the rendering process can be retained.

The second, a foveal method, interactively adjusts to preserve the fine details in the region around the viewer's gaze (which the viewer directs with a mouse) and compresses the remainder. In this instance their final tone reproduction operator is a revised version of the original Tumblin and Rushmeier [41] operator, also building on the work of Ferwerda [10] and Ward [45].

Both of these operators are straightforward in implementation and are not computationally expensive. The layering method is suited to static, synthetic scenes (displayed or printed) and the foveal method to interactive scenes (requiring a computer display).

Scheel, Stamminger and Seidel [33] developed a method that permitted tone reproduction for interactive applications by representing luminances as a texture. The luminance of each vertex is coded into texture co-ordinates, and prior to rendering these luminance co-ordinates are mapped into display luminance values through the use of Ward [45] and Ward Larson's [18] operators. This allows walkthroughs of large scenes where the tone reproduction can be adjusted frame-by-frame to the current view of the user, and focuses on tone reproduction for global illumination solutions obtained by radiosity methods. Due to interactivity, updates in tone mapping are required to account for changes in view point and viewing direction, and new factors need to be incorporated into the tone reproduction operator, such as computational speed and adaptation determination. In comparison, Tumblin et al. 's foveal method [40] was interactive to an extent, but relied on pre-computed still images where the fixation point of the viewer could change, but an interactive walkthrough was not possible.

Spatially uniform operators were chosen due to computational efficiency, and Scheel et al. based their work on operators developed by Ward [45] and Ward Larson et al. [18]. A centre-weighted average is used to determine the probability of the user's focus. The adaptation levels are computed using samples obtained through ray-tracing, and the luminance of every vertex is held in texture co-ordinates. This can then be updated frame-by-frame. This method of tone reproduction provided a new level of interactivity, but it does not take into consideration adaptation over time.

Work by **Cohen, Tchou, Hawkins and Debevec** [6] addresses the problem of HDR image display by storing and rendering high dynamic range texture maps in real time using hardware texturing architectures. In their method, HDR texture maps are stored as two separate 8-bit texture maps, one representing the high intensities and the other the low intensities. During display, these two texture maps are recombined with the aid of a dynamically adjustable exposure level to guide the overall intensity of the result.

3.2 Spatially varying operators

Oppenheim, Schafer and Stockham's [26] work on non-linear filtering in 1968 appears to be the earliest attempt at tone reproduction in computer graphics. They describe the problem of excessive dynamic range and suggest a method for simultaneously reducing dynamic range and enhancing contrast using homorphic filtering.

An image can be divided into two parts: the illumination component (the available light) and the reflection component (the ability of objects to reflect light). The illumination component, which contains large variations in luminance intensities, primarily consists of low frequencies, and the reflection component primarily consists of high frequencies. Therefore, low frequency content in an image tends to be high dynamic range, and high frequency content tends to be low dynamic range. By attenuating the low frequencies in the Fourier domain, HDR data may be compressed while the high frequencies (the low dynamic range detail) are preserved.

Further work by **Stockham** [38] in 1972 tied the concept of homorphic filtering to properties of early portions of the HVS. He developed a visual model based on these properties and used it to define a measure of image quality.

Chiu, Herf, Shirley, Swamy, Wang and Zimmerman's [5] investigation into global operators led them to believe that the solution should be local instead, as applying the same mapping to each pixel could produce incorrect results. With an HDR image there is no perfect compression curve that fits every pixel in an image, so a method of incorporating local variation is desired. They deliberately did not incorporate adaptation issues or psychophysical models into their operator; rather they experimented with a method of spatially varying image mapping. As the HVS is more sensitive to relative as opposed to absolute changes in luminance they developed a spatially non-uniform scaling function for high contrast images. Their basis was the argument that the eye is more sensitive to reflectance than luminance, so that slow spatial variation in luminance may not be greatly perceptible. The implication is that images with a wider dynamic range than the display device can be displayed without much noticeable difference if the scaling function has a low magnitude gradient. By blurring the image to remove high frequencies, and inverting the result, the original details can be reproduced, but reverse intensity gradients appear when very bright and very dark areas are in close proximity [21].

Due to the fact that it is a local operator, this model is also computationally demanding. It is also a 'hands-on' approach, based purely on experimental results and therefore does not have the advantages of the more robust, theoretical basis of other tone reproduction operators.

Schlick [34] presented practical methods of tone reproduction, concentrating on improving computational efficiency and simplifying parameters. He used a first degree rational polynomial function to map real-world luminances to display values, a function which worked well when applied uniformly to all pixels in an image. He produced three methods of mimicking local adaptation. The first of these, low pass filtering, was susceptible to halo artifacts, as was Chiu et al. 's method — a problem common among spatially varying operators. The remaining two methods did not produce as satisfactory results as the uniform approach. Nonetheless, his work is worthy in its optimisation of spatially varying techniques.

Jobson, Rahman and Woodell [15] based their method on the retinex theory [16] of colour vision, producing a multi-scale version to achieve simultaneous dynamic range compression, colour consistency and lightness rendition, testing it extensively on (real-world) test scenes and over 100 images. The retinex is a computational model of lightness and colour perception of human vision which estimates scene reflectances, and Jobson et al. modified it to perform in a functionally similar manner to human visual perception. However, in their validation they used 24-bit RGB test images where dynamic range reduction is not an issue as it can be displayed in a straightforward manner on a standard CRT. They expressed the need for refinement of their approach for images with greater maximum contrasts. Also, problems arose with scenes dominated by one colour as they violated the retinex "gray-world" assumption that the average reflectances are equal in the three spectral colour bands.

Pattanaik, Ferwerda, Fairchild and Greenberg [27] developed a technique based on a multi-scale representation of pattern, luminance, and colour processing in the HVS and addressed the problems of high dynamic range and perception of scenes at threshold and supra-threshold levels. They provided a computational model of adaptation and spatial vision for realistic tone reproduction. There are two main parts to this model: the *visual model*, which processes an input image to encode the perceived contrasts for the chromatic and achromatic channels in their band-pass mechanism; and the *display model*, which takes the encoded information and outputs a reconstructed image. Although it is computationally demanding, the model takes chromatic adaptation into account. However, as seen in other spatially varying operators, this method is susceptible to strong halo effects [39]. Although it was designed as a solution towards the tone reproduction problems of wide absolute range and high dynamic range scenes, it is a general model that can be applied across a number of areas such as image quality metrics, image compression methods and perceptually-based image synthesis algorithms [27].

Jobson, Rahman and Woodell [15] based their method on the retinex theory of colour vision, producing a multi-scale version that achieved simultaneous dynamic range compression, colour consistency and lightness rendition, and tested it extensively on (real-world) test scenes and over 100 images. The retinex is a computational model of lightness and colour perception of human vision which estimates scene reflectances, and Jobson et al. modified it to perform in a functionally similar manner to human visual perception. While this method worked well with their 24-bit RGB test images, they expressed the need for refinement of the method for images with greater maximum contrasts. Also, problems arose with scenes dominated by one colour as they violated the retinex "gray-world" assumption that the average reflectances are equal in the three spectral colour bands.

In 1999 **Tumblin and Turk** [42] produced the Low Curvature Image Simplifier (LCIS) method, a versatile technique that can accept input from synthetic sources or real-world image maps, and produces an output suitable for any display. Similar to Tumblin et al. 's [40] layering and foveal approaches, the LCIS separates the input scene into large features and fine details, compressing the former and preserving the latter. The idea stems from art where an initial sketch outlines the main structure of a picture, with details and shadings filled in later. The LCIS uses a form of anisotropic diffusion to define the fine details by scene boundaries and smooth shading. This provides a high amount of subtle detail, avoids halo artifacts, and claims moderate computational efficiency [39].

This model developed by **Pattanaik, Ferwerda, Fairchild and Greenberg** [27] was based on a multi-scale representation of pattern, luminance, and colour processing in the HVS and addressed the problems of high dynamic range and perception of scenes at threshold and supra-threshold levels. They provide a computational model of adaptation and spatial vision for realistic tone reproduction. There are two main parts: the *visual model*, which processes an input image to encode the perceived contrasts for the chromatic and achromatic channels in their band-pass mechanism; and the *display model*, which takes the encoded information and outputs a reconstructed image. Although it is still computationally demanding, the model takes chromatic adaptation into account. However, this method is susceptible to strong halo effects [39]. Although it was designed as a solution towards the tone reproduction problems of wide absolute range and high dynamic range scenes, it is a general model that can be applied across a number of areas such as image quality metrics, image compression methods and perceptually-based image synthesis algorithms [27].

Durand and Dorsey's [8] 2002 method uses an edge-preserving filter known as the bilateral filter to decompose the image into two layers — an approach which builds on Tumblin and Turk's LCIS method [42], and Tumblin et al. 's layering method [40] which it extends to photographs. A base layer (which consists of large-scale variations) is derived using bilateral filtering and contrast is reduced in this layer while visibility is preserved in the detail layer. Their final method is a faster, more robust operator that also addresses problems mentioned by Tumblin and Turk in their LCIS method [42], namely halo artifacts and diffusion at discontinuities. Again, perceptual accuracy is not the aim and their operator does not attempt to model human vision.

Recent work by **Fattal, Lischinski and Wermann** [9] presents a new method for displaying HDR scenes. Their method is computationally efficient and conceptually simple, and is based on attenuating the magnitudes of the large luminance gradients that exist in HDR scenes, compressing large gradients and preserving fine details. The changes in intensity are identified and the larger gradients are reduced and a low dynamic range image is produced. They do not make an attempt at perceptual accuracy, but instead offer an effective, fast and easy-to-use form of tone reproduction.

Reinhard, Stark, Shirley and Ferwerda's [31] recent method is analogous to photographic practice, resulting in a technique designed to suit a wide variety of images. In photography, an approach known as the Zone System is widely used. This photographic technique divides a scene into print zones ranging from pure black to white. A luminance reading is taken for a subjectively-defined middle-grey tone. Readings are taken for light and dark regions and a dynamic range can be determined, and an appropriate choice for middle-grey ensures that the maximum possible detail is retained.

Reinhard et al. take a user-specified value for middle grey. The log average luminance of the input image is then mapped to this value by linear scaling. A spatially uniform operator is then used to compress the high intensities in the image. Spatially varying tone reproduction is introduced in a manner akin to "dodging and burning" in photography. This allows contrast to be controlled locally in the image over regions bounded by large contrasts. This was based on a centre-surround function derived from a model of brightness perception by Blommaert and Martens [4]. They tested their method against existing tone reproduction operators with a broad range of HDR images.

This operator uses a slightly different definition for dynamic range. In computer graphics, dynamic range is held to be the ratio of the highest to the lowest scene luminance, whereas Reinhard et al. adopt the photographic definition that dynamic range is the ratio of the highest and lowest luminance regions where detail is visible. This results in images with ranges lower than they would be if the computer graphics definition was used. With the standard computer graphics definition of dynamic range it is difficult to know how successful compression of an HDR image will be. Using the photographic definition, Reinhard et al. correlate dynamic range with difficulty of compression, using this to predict how challenging tone reproduction for a given image will be.

This method is simple, fast and computationally efficient. As with other recent tone reproduction operators, perceptual accuracy is not attempted. Instead they aim to produce credible results.

Ashikhmin has recently produced a tone mapping operator which preserves image details and also conveys the compression of absolute brightness in a low dynamic range image using a multipass approach. First, local adaptation luminance is estimated by determining the largest sufficiently uniform neighbourhood for each pixel. Next, the tone mapping (using a TVI function — see Section 3.3) is applied, using the local adaptation information to produce a locally linear mapping. Finally, local contrast is estimated, thus preserving detail throughout the image. This approach is simple in implementation and moderate in computational expense.

3.3 Time-dependent operators

Ferwerda, Pattanaik, Shirley and Greenberg [10] developed a model which accounts for changes in colour appearance, visual acuity and temporal sensitivity while preserving global visibility. This model is based on the concept of matching JNDs for a variety of adaptation levels. It accounts for both rod and cone response (changes in colour appearance) and takes into consideration the aspect of adaptation over time.

Ferwerda et al. exploit the detectability of changes in background luminance in order to remove those frequencies imperceptible when adapted to real-world illumination. Detection threshold experiments were used as the basis of the work. This is a way of measuring visual sensitivity in a psychophysical manner. In the experiments a viewer adapts to a dark screen. On each trial a disk of light is flashed in the centre of the screen for a few hundred milliseconds. The viewer states whether or not they have seen it. If they have not, the intensity is increased on the next trial; if they have, the intensity is decreased on the next trial. This provides a detection threshold. By plotting the detection threshold against the corresponding background luminance, a threshold-versus-intensity (TVI) function is produced for both the display and the viewer.

The implementation of this model is based on Ward's 1994 operator [45]. Ward's model is used without change for cone TVI data and is extended for rod TVI data. If the level of adaptation for the real-world viewer falls in the photopic range (i.e.. above 10 cd/m^2) then a photopic tone-reproduction operator is applied (making use of the cone data), and if it falls in the scotopic range (i.e.. below 0.01 cd/m^2) then a scotopic tone reproduction operator is applied (making use of the rod data). For mesopic conditions, a photopic display luminance and a scotopic display luminance are combined appropriately.

To reproduce the loss in visual acuity, Ferwerda et al. used data from psychophysical experiments that related the detectability of square wave gratings of different spatial frequencies to changes in background luminance. Using this data it is possible to determine what spatial frequencies are visible, and thereby eradicate any extraneous data in the image. Light and dark adaptation were also considered by adding a parameter to the display luminance, the value of which changes over time.

This model is of particular importance due to the psychophysical model of adaptation that it adopts, and will prove useful for immersive display systems that cover the entire visual field so that the viewer's visual state is determined by the whole display [21].

Pattanaik, Tumblin, Yee and Greenberg [28] produced a new time-dependent tone reproduction operator to automatically create colour image sequences from any input scene. It followed the perceptual models framework proposed by Tumblin and Rushmeier with the addition of an adaptation model and appearance model to express retinal response and lightness and colour.

The adaptation model computes retina-like response signals (for rod and cone luminance and colour information) for each pixel in the scene. Using Hunt's static model of colour vision, time-dependent adaptation components are added to describe neural effects, pigment bleaching, regeneration and saturation effects. The visual appearance model assumes that the real-world viewer determines a 'reference white' and a 'reference black' and judges the appearance of any visual response against these standards. Assembling these models reproduces the appearance of scenes that evoke changes to visual adaptation. This operator is suitable for use in real-time applications as due to its spatially uniform model of adaptation it does not require extensive processing.

Durand and Dorsey [7] presented an interactive tone mapping model which made use of visual adaptation knowledge. They also proposed extensions to Ferwerda et al. 's tone mapping operator and incorporated it into a model for the display of global illumination solutions and interactive walkthroughs. This model involves time-dependent tone mapping and light adaptation, and extends Ferwerda et al. 's by including a blue-shift for viewing night scenes and by adding chromatic adaptation. For the interactive implementation, work by Tumblin et al. [40] and Scheel et al. [33] was used to take advantage of the observer's gaze, allowing a weighted average to be used. Photographic exposure metering used in photography is employed to better calculate the adaptation level. Loss of visual acuity is simulated in the same manner of Ferwerda et al. by use of a 2D Gaussian blur filter. The scene is rendered as normal, with interactivity introduced by tone mapping computed on the fly, accelerated by caching the function in look-up tables.

3.4 Related effects

Replication of visual effects that are related to the area of tone reproduction include the modelling of glare. **Spencer, Shirley, Zimmerman and Greenberg** [36] developed a method for replicating glare effects. The idea of adding glare effects was previously recognised by Nakamae et al.[23], although their algorithm did not account for the visual masking effects of glare.

Spencer et al. produced psychophysically-based algorithms for adding glare to digital images, simulating the flare and bloom seen around very bright objects, and carried out a psychophysical test to demonstrate that these effects increased the apparent brightness of a light source in an image. While highly effective, glare simulation is computationally expensive.

Algorithm	Spatially Uniform	Spatially Varying	Time dependent	Attributes
Oppenheim et al. 1968 [26]		√		Attenuation of low frequency (HDR) data.
Stockham 1972 [38]		√		Extends Oppenheim et al. using a visual model.
Miller et al. 1984 [22]	√			Uses perceptual data.
Upstill 1985 [43]	√			Uses perceptual model.
Tumblin and Rushmeier 1993 [41]	√			Uses perceptual model. Preserves brightness. Does not preserve visibility or account for adaptation. Grey scale only.
Chiu et al 1993 [5]		√		Preserves local contrast. Ad hoc. Computationally demanding.
Ward 1994 [45]	√			Uses perceptual model. Preserves contrast - crucial for predictive lighting analysis. Clipping of very high and very low values. Does not consider complexities of typical workplace viewing.
Schlick 1994 [34]		√		Speed and simplification of uniform and varying operators.
Ferwerda et al. 1996 [10]	√		√	Uses perceptual model. Accounts for changes in threshold visibility, colour appearance, visual acuity and sensitivity over time. Useful for immersive displays.
Ward Larson et al. 1997 [18]	√			Uses perceptual model. Histogram adjustment. Preserves local contrast visibility. Uses models for glare, colour sensitivity and visual acuity to increase perceptual realism.
Jobson et al. 1997 [15]		√		Multi-scale retinex model - perceptually valid. Problems with monochrome scenes and maximum contrasts outside 24-bit RGB range.
Pattanaik et al. 1998 [27]		√		Multi-scale psychophysical representation of pattern, luminance and colour processing resulting in increased perceptual fidelity.
Tumblin et al. 1999 [40]	√			Uses perceptual model. Layering method for static, synthetic images. Foveal method for interactive scenes.
Tumblin and Turk 1999 [42]		√		LCIS method preserves subtle details.
Pattanaik et al. 2000 [28]	√		√	Psychophysical operator with time-dependent adaptation and appearance models.
Scheel et al. 2000 [33]	√			Interactive method using texturing hardware.
Durand and Dorsey 2000 [7]	√		√	Interactive method with time-dependent adaptation and simulation of visual acuity and chromatic adaptation.
Cohen et al. 2001 [6]	√			Real-time method from HDR texture maps using graphics hardware.
Durand and Dorsey 2002 [8]		√		Uses edge-preserving filtering to decompose image, reduce contrast and preserve details.
Fattal et al. 2002 [9]		√		Computationally efficient and simple operator that attenuates the magnitudes of large luminance gradients.
Reinhard et al. 2002 [31]		√		Method analogous to photographic technique. Suits a wide variety of images; fast and computationally efficient.
Ashikhmin 2002 [2]		√		Uses simple functional perceptual model to preserve image details and absolute brightness information.

Table 1: Tone Reproduction Operators: summary of attributes

4 Conclusion

Until advances in hardware provide us with a more advanced form of display we will have to depend on tone reproduction operators to deliver the desired perceptual effect.

Evolution of display technology has seen the beginning of a move away from the standard CRT monitor to flat screen LCD displays and micro-mirror projection systems, but these have still to become commonplace, and have disadvantages of their own. In the case of LCD displays, limitations are imposed due to shortcomings involving angular dependence, temperature, channel constancy, resolution and a lack of control of gamma.

The development of a high dynamic range viewer will allow for testing of existing tone reproduction operators, allowing us to apply the most effective models for our purpose. (Greg Ward has produced an experimental high contrast stereoscopic display with a contrast ratio of 5000:1 [44].) Although our display capabilities are limited, it is important to ensure that the information is stored in a relevant device-independent representation so that none of the HDR information is lost, thus preserving display options. Formats such as the SGI LogLuv TIFF, which can hold 38 orders of magnitude in its 32-bit version, have been recommended [17, 46].

Knowledge of the HVS is still limited, and modelling its characteristics for the purpose of perceptually accurate tone mapping will be complex and time consuming. The lack of comprehensive image metrics in graphics also limits the study. At present, the answer is to use the most appropriate method for the situation. Depending on requirements, a number of different operators are available for use and they must be selected on the premise of the 'best tool for the job'. There is undoubtedly a need for the validation of tone reproduction operators, preferably through psychophysical comparison.

Acknowledgements

Thanks to those who gave their time to proofread and discuss sections of this paper.

References

[1] I. Ashdown and P. Franck. Luminance gradients: Photometric analysis and perceptual reproduction. In *IESNA Annual Conference Technical Papers*. Illuminating Engineering Society of North America, 1995.

[2] M. Ashikhmin. A tone mapping algorithm for high contrast images. In *13th Eurographics Workshop on Rendering*. Eurographics, June 2002.

[3] H. R. Blackwell. *An Analytical Model for Describing the Influence of Lighting Parameters upon Visual Performance*, volume 1: Technical Foundations. Commission Internationale De L'Eclairage, 1981.

[4] F. J. J. Blommaert and J. B. Martens. An object-oriented model for brightness perception. *Spatial Vision*, 5(1):15–41, 1990.

[5] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Graphics Interface '93*, pages 245–253, Toronto, Ontario, Canada, May 1993. Canadian Information Processing Society.

[6] Jonathan Cohen, Chris Tchou, Tim Hawkins, and Paul Debevec. Real-Time high dynamic range texture mapping. In *12th Eurographics Workshop on Rendering*, pages 313–320. Eurographics, June 2002.

[7] Fredo Durand and Julie Dorsey. Interactive tone mapping. In *Rendering Techniques 2000: 11th Eurographics Workshop on Rendering*, pages 219–230. Eurographics, June 2000. ISBN 3-211-83535-0.

[8] Frédo Durand and Julie Dorsey. Fast bilateral filtering for the display of high dynamic range image. In John Hughes, editor, *SIGGRAPH 2002 Conference Graphics Proceedings*, Annual Conference Series, pages 257–265. ACM Press/ACM SIGGRAPH, 2002.

[9] Raanan Fattal, Dani Lischinski, and Micheal Werman. Gradient domain high dynamic range compression. In *Proceedings of ACM SIGGRAPH 2002*, Computer Graphics Proceedings, Annual Conference Series. ACM Press / ACM SIGGRAPH, July 2002.

[10] James A. Ferwerda, Sumanta Pattanaik, Peter S. Shirley, and Donald P. Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of SIGGRAPH 96*, Computer Graphics Proceedings, Annual Conference Series, pages 249–258, New Orleans, Louisiana, August 1996. ACM SIGGRAPH / Addison Wesley. ISBN 0-201-94800-1.

[11] James A. Ferwerda, Sumanta N. Pattanaik, Peter S. Shirley, and Donald P. Greenberg. A model of visual masking for computer graphics. In *Proceedings of SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, pages 143–152, Los Angeles, California, August 1997. ACM SIGGRAPH / Addison Wesley. ISBN 0-89791-896-7.

[12] S. Gibson. *Efficient Radiosity Simulation using Perceptual Metrics and Parallel Processing*. Phd thesis, University of Manchester, September 1998.

[13] Donald P. Greenberg, Kenneth E. Torrance, Peter S. Shirley, James R. Arvo, James A. Ferwerda, Sumanta Pattanaik, Eric P. F. Lafortune, Bruce Walter, Sing-Choong Foo, and Ben Trumbore. A framework for realistic image synthesis. In *Proceedings of SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, pages 477–494, Los Angeles, California, August 1997. ACM SIGGRAPH / Addison Wesley. ISBN 0-89791-896-7.

[14] R. W. G. Hunt. *The Reproduction of Colour in Photography, Printing and Television*. Fountain Press, Tolworth, 5th edition, 1995.

[15] D. J. Jobson, Z. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6(7):965–976, July 1997.

[16] E. H. Land and J. J. McCann. Lightness and the retinex theory. *Journal of the Optical Society of America*, 61(1):1–11, 1971.

[17] Gregory Ward Larson. Logluv encoding for full-gamut, high-dynamic range images. *Journal of Graphics Tools*, 3(1):15–31, 1998. ISSN 1086-7651.

[18] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, October - December 1997. ISSN 1077-2626.

[19] K. Matkovic, L. Neumann, and W. Purgathofer. A survey of tone mapping techniques. In *13th Spring Conference on Computer Graphics*, pages 163–170, 1997.

[20] Kresimir Matkovic and Lazlo Neumann. Interactive calibration of the mapping of global illumination values to display devices. In *Proceedings of the Twelfth Spring Conference on Computer Graphics*, Comenius University, Bratslavia, Slovakia, June 1996. Available from http://cg.tuwien.ac.at/ wp/SCCG96-proceedings.

[21] Ann McNamara. Visual perception in realistic image synthesis. *Computer Graphics Forum*, 20(4):211–224, 2001. ISSN 1067-7055.

[22] N.J. Miller, P.Y. Ngai, and D.D. Miller. The application of computer graphics in lighting design. *Journal of the IES*, 14:6–26, 1984.

[23] Eihachiro Nakamae, Kazufumi Kaneda, Takashi Okamoto, and Tomoyuki Nishita. A lighting model aiming at drive simulators. In Forest Baskett, editor, *Computer Graphics (SIGGRAPH '90 Proceedings)*, volume 24, pages 395–404, August 1990.

[24] L. Neumann, K. Matkovic, and W. Purgathofer. Automatic exposure in computer graphics based on the minimum information loss principle. In *Computer Graphics International 1998*, Hanover, Germany, June 1998. IEEE Computer Society.

[25] László Neumann, Kresmir Matkovic, Attila Neumann, and Werner Purgathofer. Incident light metering in computer graphics. *Computer Graphics Forum*, 17(4):235–247, 1998. ISSN 1067-7055.

[26] A. Oppenheim, R. Schafer, and T. Stockham. Nonlinear filtering of multiplied and convolved signals. In *Proceedings of the IEEE*, volume 56, pages 1264–1291, August 1968.

[27] Sumanta N. Pattanaik, James A. Ferwerda, Mark D. Fairchild, and Donald P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of SIGGRAPH 98*, Computer Graphics Proceedings, Annual Conference Series, pages 287–298, Orlando, Florida, July 1998. ACM SIGGRAPH / Addison Wesley. ISBN 0-89791-999-8.

[28] Sumanta N. Pattanaik, Jack E. Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for realistic image display. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 47–54. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, July 2000. ISBN 1-58113-208-5.

[29] C. Poynton. Frequently asked questions about gamma. http//:www.inforamp.net/poynton/.

[30] C. Poynton. *A Technical Introduction to Digital Video*. John Wiley and Sons, 1996.

[31] Erik Reinhard, Michael Stark, Peter Shirley, and Jim Ferwerda. Photographic tone reproduction for digital images. In *Proceedings of ACM SIGGRAPH 2002*, Computer Graphics Proceedings, Annual Conference Series. ACM Press / ACM SIGGRAPH, July 2002.

[32] G. Rougeron and B. Péroche. Color fidelity in computer graphics: A survey. *Computer Graphics Forum*, 17(1):3–16, 1998. ISSN 1067-7055.

[33] A. Scheel, M. Stamminger, and Hans-Peter Seidel. Tone reproduction for interactive walkthroughs. *Computer Graphics Forum*, 19(3):301–312, August 2000. ISSN 1067-7055.

[34] C. Schlick. Quantization techniques for visualization of high dynamic range pictures. In *5th Eurographics Workshop on Rendering*. Eurographics, June 1994.

[35] Christophe Schlick. High dynamic range pixels. In *Graphics Gems IV*, pages 422–429. Academic Press, Boston, 1994. ISBN 0-12-336155-9.

[36] Greg Spencer, Peter S. Shirley, Kurt Zimmerman, and Donald P. Greenberg. Physically-based glare effects for digital images. In *Proceedings of SIGGRAPH 95*, Computer Graphics Proceedings, Annual Conference Series, pages 325–334, Los Angeles, California, August 1995. ACM SIGGRAPH / Addison Wesley. ISBN 0-201-84776-0.

[37] S. S. Stevens and J. C. Stevens. Brightness function: Parametric effects of adaptation and contrast. *Journal of the Optical Society of America*, 53(1139), 1960.

[38] T. Stockham. Image processing in the context of a visual model. In *Proceedings of the IEEE*, volume 60, pages 828–842, 1972.

[39] J. Tumblin. *Three Methods of Detail-Preserving Contrast Reduction for Displayed Images*. Phd thesis, Georgia Institute of Technology, December 1999.

[40] Jack Tumblin, Jessica K. Hodgins, and Brian K. Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, January 1999. ISSN 0730-0301.

[41] Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics & Applications*, 13(6):42–48, November 1993.

[42] Jack Tumblin and Greg Turk. Lcis: A boundary hierarchy for detail-preserving contrast reduction. In *Proceedings of SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, pages 83–90, Los Angeles, California, August 1999. ACM SIGGRAPH / Addison Wesley Longman. ISBN 0-20148-560-5.

[43] S.D. Upstill. *The Realistic Presentation of Synthetic Images*. PhD thesis, Computer Science Division, University of California, Berkeley, 1985.

[44] G. Ward. A wide field, high dynamic range, stereographic viewer. In *Proceedings of PICS 2002*, Portland, Oregon, 2002.

[45] Greg Ward. A contrast-based scalefactor for luminance display. In *Graphics Gems IV*, pages 415–421. Academic Press, Boston, 1994. ISBN 0-12-336155-9.

[46] Gregory Ward. High dynamic range imaging. In *Proceedings of the Ninth Colour Imaging Conference*, November 2001.

[47] Hector Yee, Sumanta Pattanaik, and Donald P. Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics*, 20(1):39–65, January 2001. ISSN 0730-0301.