

Lightness Perception in Tone Reproduction for High Dynamic Range Images

Grzegorz Krawczyk[†] and Karol Myszkowski[‡] and Hans-Peter Seidel[§]

MPI Informatik, Saarbrücken, Germany

Abstract

An anchoring theory of lightness perception comprehensively explains many characteristics of human visual system such as lightness constancy and its spectacular failures which are important in the perception of images. We present a novel approach to tone mapping of high dynamic range (HDR) images which is inspired by the anchoring theory. The key concept of this method is the decomposition of an HDR image into areas (frameworks) of consistent luminance and the local calculation of the lightness values. The net lightness of an image is calculated via the merging of the frameworks proportionally to their strength. We stress out the importance of relating the luminance to a known brightness value (anchoring) and investigate the advantages of anchoring to the luminance value perceived as white. We validate the accuracy of the lightness reproduction in the presented algorithm by simulating a well known perception experiment. Our approach does not affect the local contrast and preserves the natural colors of an HDR image due to the linear handling of luminance.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Display algorithms

1. Introduction

Lightness is a perceptual quantity measured by the human visual system (HVS) which describes the amount of light reflected from the surface normalized for the illumination level. Contrary to brightness, which describes a visual sensation according to which an area exhibits more or less light, the lightness of a surface is judged relative to the brightness of a similarly illuminated area that appears to be white. Lightness constancy is an important characteristics of the HVS which leads to a similar appearance of the perceived objects independently of the lighting and viewing conditions [Pal99]. While observing the images presented on display devices, it would be desirable to reproduce the lightness perception corresponding to the observation conditions in the real world. This is not an easy task because the lightness constancy achieved by the HVS is not perfect and many of its failures appear in specific illumination conditions or

even due to changes in the background over which an observed object is imposed [Gil88]. It is well known that the lightness constancy increases for scene regions that are projected over wider retinal regions [Roc83]. This effect is reinforced for objects whose perceived size is larger even for the same retinal size [GC94]. The reproduction of images on display devices introduces further constraints in terms of a narrower field of view and limitations in the luminance dynamic range. Some failures of lightness constancy still appear in such conditions (simultaneous contrast for instance), but other, such as the Gelb illusion, cannot be observed on a display device.

For images that capture real world luminance levels, the so-called high dynamic range (HDR) images, the problem of a correct lightness level reproduction must be addressed. Fortunately, such a reproduction is feasible because the HVS has a limited capacity to detect the differences in the absolute luminance levels and it concentrates more on the aspects of spatial patterns when comparing two images [Pal99]. However, a successful lightness reproduction algorithm should properly account for the lightness constancy failures that cannot be evoked in the HVS using a display device. Clearly,

[†] e-mail: krawczyk@mpi-inf.mpg.de

[‡] e-mail: karol@mpi-inf.mpg.de

[§] e-mail: hpseidel@mpi-inf.mpg.de

we need to embed a lightness perception model into the processing pipeline of HDR images in order to improve the fidelity of their display.

The problem of lightness perception has been studied extensively in the last two centuries (refer to [Pal99] for a detailed historical account). The most prominent theories follow Wallach's observation that the perceived lightness depends on the ratio of the luminance at edges between neighboring image regions. In the retinex theory [LM71] it is assumed that even for remote image regions such a ratio can be determined through the edge integration of luminance ratios along an arbitrary path connecting those regions. Lightness constancy can be well modeled by the retinex algorithm under the condition that the illumination changes slowly, which effectively means that sharp shadow borders cannot be properly processed. To overcome this problem, Gilchrist and his collaborators suggested that the HVS performs an edge classification to distinguish illumination and reflectance edges [Gil77]. This led to the concept of the decomposition of retinal images into the so called intrinsic images [BT78, Are94] with reflection, illumination, depth and other information stored in independent image layers.

Modern lightness perception theories based on intrinsic images deal with lightness constancy very successfully, however they have problems with the modeling of apparent failures of lightness constancy [GKB*99]. Also, while they provide relative lightness values for various scene regions, they fail to assign corresponding absolute lightness values for given observation conditions. In fact, it is enough to find only one such a corresponding absolute value (the so called *anchor* value) and the remaining values immediately can be found through the known ratios. The problem of lightness constancy failures and absolute lightness assignment is addressed by an *anchoring theory* of lightness perception developed by Gilchrist et al. [GKB*99] which is supported by extensive experimental studies with the human subjects.

In this paper we investigate the application of the anchoring theory [GKB*99] to the tone mapping problem, which deals with the rendering of HDR images on low dynamic range (LDR) display devices. The overall goal of our research is to address the problem of the correct lightness reproduction during the dynamic range compression, limiting the distortions of the local contrasts and the change in the appearance of the original colors whenever possible. During the tone mapping, apart from the assignment of absolute lightness values to various image regions, as predicted by the anchoring theory, we may also need to scale the resulting lightness ratios to adapt them to the limited dynamic range. Since we deal with real world scenes, we introduce a new approach to analyze an HDR image in terms of areas that have common properties, called the frameworks. On the technical level, this requires an algorithm for automatic extraction of such frameworks from complex HDR images. Finally, we stress out the importance of relating the luminance values

to a known brightness level – anchoring. We investigate the advantages of anchoring to luminance value perceived as white, instead of middle-gray which is a common practice [FPSG96, THG99, PTYG00, RSSF02]. We then simulate a perception experiment to illustrate the accuracy of the lightness reproduction in the presented tone mapping algorithm.

The paper is organized as follows. Section 2 provides an overview of existing tone mapping operators from the standpoint of lightness perception. In Section 3 we briefly overview the anchoring theory of lightness perception. In Section 4 we demonstrate our application of this theory to the tone mapping problem, which involves the image decomposition into frameworks and the computation of anchor points for those frameworks. In Section 5 we present tone mapped images obtained using our technique and we provide information concerning its performance. We conclude the paper in Section 6 and we propose directions of future research.

2. Previous Work

The problem of tone mapping has been extensively studied in computer graphics for over a decade (refer to [DCWP02] for a survey). In this section, we analyze the existing algorithms in view of their ties to lightness perception theories.

A number of tone mapping operators based on models of the HVS have been proposed. Tumblin and Rushmeier [TR93] used a model of brightness perception that was based on the power-law relationship between the brightness and the corresponding luminance, as proposed by Stevens and Stevens [SS60]. The main objective was to preserve a constant relationship between the brightness of a scene perceived on a display and its real counterpart for any lighting condition. Other operators included threshold models of the contrast perception for the scene luminance compression [War94, FPSG96, WLRP97, PFFG98, Ash02]. To account for the adaptation to various luminance levels the threshold versus intensity function has been used. Pattanaik et al. [PTYG00] and Reinhard et al. [RD05] used sigmoid functions which modeled the retinal response of cones and rods for the luminance compression.

There were some attempts of a direct application of the retinex theory [LM71] to tone mapping. Jobson et al. [JRW97] proposed a multi-resolution retinex algorithm for luminance compression, which unfortunately lead to halo artifacts for the HDR images along high contrast edges. Inspired by the lightness perception model, developed by Horn [Hor74], Fattal et al. [FLW02] proposed a successful gradient domain tone mapping operator. Fattal et al. observed that any large contrast in the image must give rise to large magnitude luminance gradients, while textures and other fine details result in gradients of much smaller magnitude. Their algorithm identified such large gradients and attenuated them without altering their directions, which lead to halo-free images of a good quality.

A concept of intrinsic images [BT78, Are94], separating the illumination and reflectance (details) layers, inspired many tone mapping operators. The high contrast of the illumination layer is usually reduced by scaling, while the details layer (assumed to be of low contrast) is preserved. The idea was first introduced by Tumblin et al. [THG99], who assumed that these layers are explicitly provided, which is the case only for synthetic images. Later, several methods for an automatic layer separation have been introduced. The LCIS operator [TT99] separates the image into large scale features (presumably illumination) and fine details. A much better separation has been achieved using the bilateral filter [DD02].

All of the aforementioned methods to a certain extent take into account the findings in psychophysics and physical processing in the retina, but none is explicitly influenced by the theoretical research in lightness perception. The tone mapping based on the separation of the HDR image into illumination and detail layers closely resemble the intrinsic image models [BT78, Are94]. These models are known as one of the most advanced models of lightness theory, however they do not define how to treat the luminance within each layer. In general, the anchoring in tone mapping is often either neglected or applied only indirectly via the luminance normalization by the logarithmic average of an HDR image.

3. Anchoring Theory Of Lightness Perception

The recently presented model of an anchoring theory of lightness perception by Gilchrist et al. [GKB*99] provides an unprecedented account of empirical experiments for which it provides a sound explanation. The theory is qualitatively different from the intrinsic image models and is based on a combination of global and local anchoring of lightness values. We introduce the main concepts of this theory in the following sections. We first discuss the rules of anchoring in simple images and later explain how to apply them to complex scenes.

3.1. Anchoring Rule

In order to relate the luminance values to lightness, it is necessary to define at least one mapping between the luminance value and the value on the scale of perceived gray shades – *the anchor*. Once such an anchor is defined, the lightness value for each luminance value can be estimated by the luminance ratio between the value and the anchor. This estimation is referred to as *scaling*. As noted before, the anchor cannot be defined once for the absolute luminance values, but rather must be tied to some measure of relative luminance values. Practically, two different approaches to anchoring are known: the average luminance rule and the highest luminance rule.

The average luminance rule derives from the adaptation-level theory [Hel64] and states that the average lumi-

nance in the visual field is perceived as middle gray. Thus the relative luminance values should be anchored by their average value to middle gray. This assumption was later commonly adopted in tone mapping techniques [FPSG96, THG99, PTYG00, RSSF02].

Initially the highest luminance rule defined the anchor as a mapping of the highest luminance in the visual field to a lightness value perceived as white. However, the perception of self-luminous surfaces contradicts this rule. When for instance a relatively small white disc is surrounded by a large dark area (an increment test stimuli), the white disc appears self-luminous, i.e. produces the impression of being lighter than white. Apparently the perception of lightness is affected by a relative area [LG99]. There is a tendency of the highest luminance to appear white and a tendency of the largest area to appear white. Therefore the highest luminance rule was redefined based on this experimental evidence. As long as there is no conflict, i.e. the highest luminance covers the largest area, the highest luminance becomes a stable anchor. However, when the darker area becomes larger, the highest luminance starts to be perceived as self-luminous. The anchor becomes a weighted average of the luminance proportionally to the occupying area.

The experimental evaluation of the average luminance rule versus the highest luminance rule was presented in [LG99]. In this study, the visual field of the observers was limited to a large acrylic hemisphere whose one half was painted matte black and the other half was painted middle-gray. The experiment was conducted in the isolated conditions to prevent the uncontrolled influence of other stimuli. Li and Gilchrist reported that the middle-gray half was seen by the observers as fully white, while the black half was seen as dark gray. Additionally, when the black area was increased and became considerably larger than the middle-gray area, the perceptual effect of self-luminosity for the middle-gray part was reported. Other findings, based on Mondrians [Pal99], which are more complex stimuli, agree with these conclusions [GC94]. Rich experimental evidence decisively favors the highest luminance rule over the average luminance rule.

3.2. Complex Images

The anchoring rule, described in the previous section, cannot be applied directly to complex images in an obvious way. Instead, Gilchrist et al. [GKB*99] introduce the concept of decomposition of an image into components, *frameworks*, in which the anchoring rule can be applied directly. In the described theory, following the *gestalt* theorists, frameworks are defined by regions of common illumination. For instance, all objects being under the same shadow would constitute a framework. A real-world image is usually composed of multiple frameworks.

The framework regions can be organized in an adjacent

or a hierarchical way and their areas may overlap. The lightness of a target is computed according to the anchoring rule in each framework. However, if a target in a complex image belongs to more than one framework, it may have different lightness values when anchored within different frameworks. According to the model, the net lightness of a given target is predicted as a weighted average of its lightness values in each of the frameworks in proportion to the *articulation* of this framework. The articulation of a framework is determined by the variety of luminance values it contains in such a way that frameworks with low variance have less influence on the net lightness.

4. Tone Mapping Method

Based on the lightness perception theory discussed in the previous section we derive a tone mapping algorithm for contrast reduction in HDR images. The algorithm takes as an input an HDR image defined by floating point RGB values that are linearly related to luminance, and produces a displayable LDR image as a result. The contrast reduction process is solely based on the luminance channel.

We first decompose the input scene into overlapping frameworks. Each pixel of the HDR image is described by the probability of its belongingness to each framework. Next, we estimate the anchor in each framework, i.e. the luminance value perceived as white. We then compute the local pixel lightness within each framework. Finally, we calculate the net lightness of each pixel by merging the individual frameworks into one image proportionally to the pixel probability values. The result is suitable to be viewed on an LDR display device.

In the following sections we explain in details each stage of the presented tone mapping algorithm and provide the implementation details.

4.1. Decomposition into Frameworks

We decompose an HDR image into frameworks based on the pixel intensity value. As we later show in practice, the contrast range typical to everyday situations is wide enough to allow us to identify in the image for example the areas of a dark shadow on a sunny day, the dim interior of a room with a window view on a sunny outdoor, the street light illumination in a night scene, and so on.

To find a plausible decomposition into frameworks, we experimented with several segmentation algorithms and noted that the mean shift segmentation [CM02] produced the most appropriate results. However, segmentation algorithms strictly assign each pixel to only one segment, therefore in our application pixels were often incorrectly segmented in the areas where they should belong to multiple frameworks. Furthermore, using pure segmentation methods, we were unable to correctly define smooth transitions between the borders of frameworks. We therefore decided to tailor a custom

decomposition method. We explain our algorithm on an example HDR image (Figure 1) using Figures 2 and 3 as a reference.

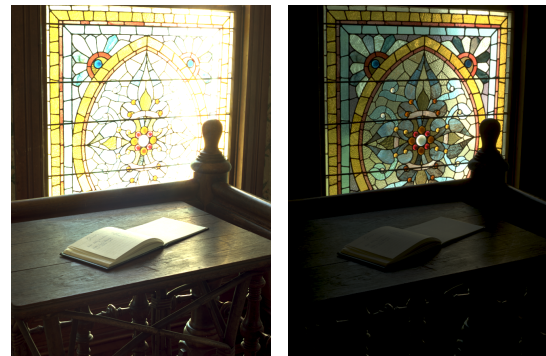


Figure 1: DESK – an example HDR image from the *OpenEXR* samples. To give an impression of the dynamic range in the scene, the left image is exposed for the dim interior and the right image is exposed for the stained glass.

We represent a framework as a probability map over the whole image in which a probability of belonging to this framework is assigned to each pixel. We define an area of a framework as a group of pixels of the HDR image whose probability of belonging to this framework is above 0.6. A valid framework must have a non-zero area. For the purpose of tone mapping for LDR displays we impose a further constraint on the dynamic range in the framework's area, which cannot exceed two orders of magnitude.

To obtain a decomposition that meets the aforementioned requirements, we implement the following method. We start with the standard K-means clustering algorithm to find the centroids that provide an appropriate segmentation of the HDR image into frameworks. We operate on a histogram in the \log_{10} of luminance. We define a constraint on the difference between two neighboring centroids, which should not be below one, to prevent frameworks from representing too similar illumination. We initialize the K-means algorithm with values ranging from the minimum to maximum luminance in the HDR image with a luminance step equal to one order of magnitude and we execute the iterations until the algorithm converges. An example histogram of the HDR image with converged centroids is shown in Figure 2 (top histogram).

Often, the converged centroids represent empty segments or may require refinement to meet the imposed constraints. First, we remove the centroids to which no image pixels were grouped. Then, we iteratively merge centroids if the difference between them is below one. In each iteration, the two closest centroids are merged together and the new centroid value is equal to their weighted average proportional to their

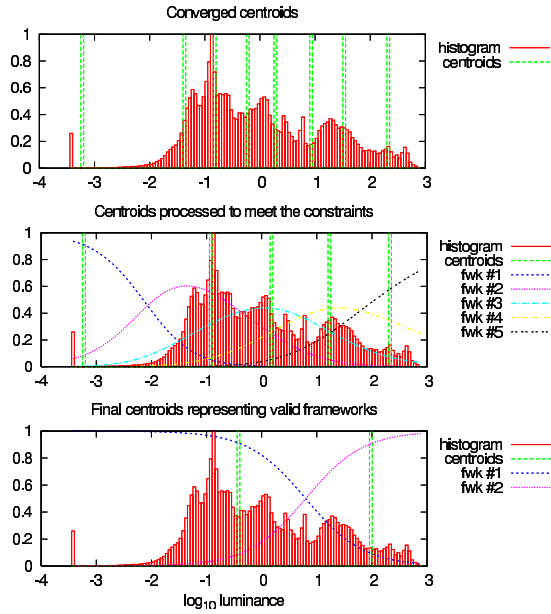


Figure 2: The histogram of the HDR image from Figure 1 illustrating the estimation of centroids which provide an appropriate decomposition into frameworks. In the middle and bottom histograms the belongingness functions are shown for each framework. The maxima of the belongingness functions do not always match the centroids due to the normalization.

area:

$$C_{i,j} = \frac{C_i \cdot S_i + C_j \cdot S_j}{S_i + S_j} \quad (1)$$

where C_i and C_j are the values of the too close centroids, and S_i and S_j denote the number of pixels clustered to this centroids. In the middle histogram in Figure 2 we show processed centroids which meet our constraints.

Given the centroid values, we initially assign the probability values based on the difference between the pixel value and the centroid. We model such a *belongness* to the centroid with a Gaussian function:

$$P_i(x, y) = e^{-\frac{(C_i - Y(x,y))^2}{2\sigma^2}} \quad (2)$$

where P_i represents the probability map for framework i , C_i is the centroid for that framework, Y denotes the luminance of the HDR image (both C_i and Y are in the \log_{10} space), and the variance σ equals to a maximum distance between adjacent centroids. The belongingness values are normalized to correctly represent the probabilities. At this stage, some centroids may still represent a framework with an empty area (according to our definition). We iteratively remove these centroids by merging them to the closest neighboring centroid using equation (1) and recalculate the probabilities.

In Figure 2 in the middle histogram, we show the modelled probabilities for each centroid. Apparently, several frameworks do not contain any pixels with a probability above 0.6 and therefore should be removed. The bottom histogram of in the same figure illustrates the final centroids with probability functions that define valid frameworks.

Next, we spatially process the probability map of each framework. The goal of spatial processing is to smooth local small variations in the probability values which may appear due to textures. At the same time, however, it is important to preserve high variations which could appear on the borders of frameworks where high illumination contrast appears. The bilateral filter [TM98] is an appropriate image processing tool for this purpose. We filter the probability map of each framework with a bilateral filter in which the range variance is set to 0.4 and the spatial variance to the half of the smaller dimension of the HDR image.

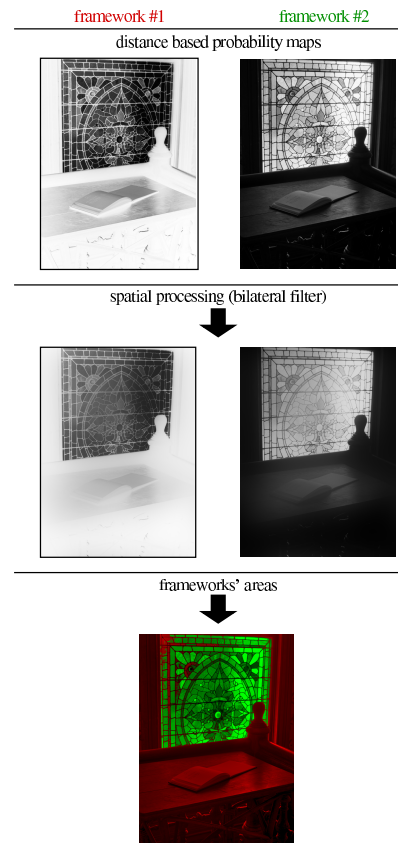


Figure 3: Processing of probability maps which represent the decomposition into frameworks. The probability range 0 : 1 is linearly mapped to a gray scale, therefore white corresponds to the highest probability and black to the lowest. In the frameworks' area, red and green colors represent the interior framework and the stained glass framework respectively.

In Figure 3 we show the probability maps describing the decomposition into frameworks of an example HDR image. Using our methods, we can successfully identify the dim interior and the brightly back lit stained glass. Some parts of the stained glass in the framework's area image may appear to be assigned incorrectly, however the underlying probability values are close to 0.5 what assures that no artifacts in the further processing will be introduced there.

4.2. Articulation of Framework

Apart from the illumination conditions, the probability of a pixel belonging to a particular framework also depends on the articulation of the framework (see Section 3.2). If a framework is highly articulated, the pixels are strongly anchored within this framework even if their probability of belonging to this framework is not high.

For the purpose of tone mapping, we estimate the articulation independently for each framework based on its dynamic range. A framework whose dynamic range is higher than one order of magnitude has a maximum articulation and as the dynamic range goes down to zero the articulation reaches the minimum. We model the amount of articulation using a Gaussian function:

$$A_i = 1 - e^{-\frac{(\max Y_i - \min Y_i)^2}{2 \cdot 0.33^2}} \quad (3)$$

where A_i denotes the articulation of the framework i , and $\min Y_i$ and $\max Y_i$ represent the minimum and maximum \log_{10} luminance in the area of this framework. The plot of the articulation function is shown in Figure 4.

We apply the articulation factor to frameworks by multiplying their probability maps P_i by their respective articulation factors A_i . We then normalize the probability maps and as such obtain the final result of the decomposition into frameworks.

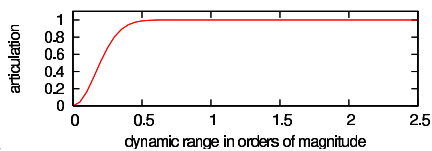


Figure 4: Plot of the articulation factor based on the dynamic range of a framework.

Usually, all frameworks in an image will have a maximum articulation. Sometimes however, a uniform area like a background may constitute a framework due to its unique illumination. Articulation prevents such a background framework to play an important role in the computation of the net lightness by minimizing the local anchoring of pixels to this framework in favor of other frameworks. In an extreme situation, when all frameworks have minimum articulation, the framework with the highest anchor is assigned a maximum articulation, thus imposing the global anchoring in the image.

4.3. Estimation of Anchor

Having the HDR image decomposed into frameworks, we estimate an anchor within each framework. Since we employ the highest luminance rule, we need to find the luminance value that would be perceived as white, in case a given framework would be observed as stand-alone.

As we discussed in Section 3.1, although we apply the highest luminance rule, we cannot directly use the highest luminance in the framework as an anchor. Seemingly, there is a relation between what is locally perceived as white and its area. If the highest luminance covers a large area it becomes a stable anchor. However, if the highest luminance is largely surrounded by darker pixels, these pixels have a tendency to appear white and the highest luminance appears as self-luminous. This implies an area related approach to the estimation of the local anchor. Therefore, we estimate the local anchor by removing 5% of all pixels in the framework's area that have the highest luminance and then take the highest luminance of the rest of the pixels as the anchor (i.e., we compute the 95-th percentile). With this approach we are able to skip potential pixels that represent self-luminous areas in the scene as for instance highlights. In case if there are no highlights, the anchor is only slightly underestimated.

On the other hand, since the 5% is an empirical value, computing the 95-th percentile may appear to be an unstable approach. Therefore we performed a stability test. We checked whether we obtain the same anchor in all frameworks for the HDR images of the same scene, but at three different resolutions – 100%, 60% and 30% of the original resolution. In this test we included all the images presented in the results section and a total of 50 other images from our database. The maximum difference between the anchors did not exceed 0.05 on a \log_{10} scale, which is a stable result for our purposes. It is important to note, that the rescaling of an image affects its dynamic range what may have a direct influence on the anchor.

In Figure 5 we show the two frameworks identified in the example HDR image with their lightness computed according to the local anchor. For instance, the luminance of the white paper in the open book is accurately estimated as the local anchor in the framework #1.

4.4. Merging Frameworks

In the final stage of the tone mapping process, we compute the global lightness of the pixels by merging the frameworks. We individually process each framework in turn. We shift the original luminance values of the HDR image according to the locally estimated lightness value and proportionally to the probability map:

$$L(x,y) = Y(x,y) - \sum_i W_i \cdot P_i(x,y) \quad (4)$$

where L denotes the final lightness value, Y the original luminance of the HDR image, W_i the local anchor of frame-

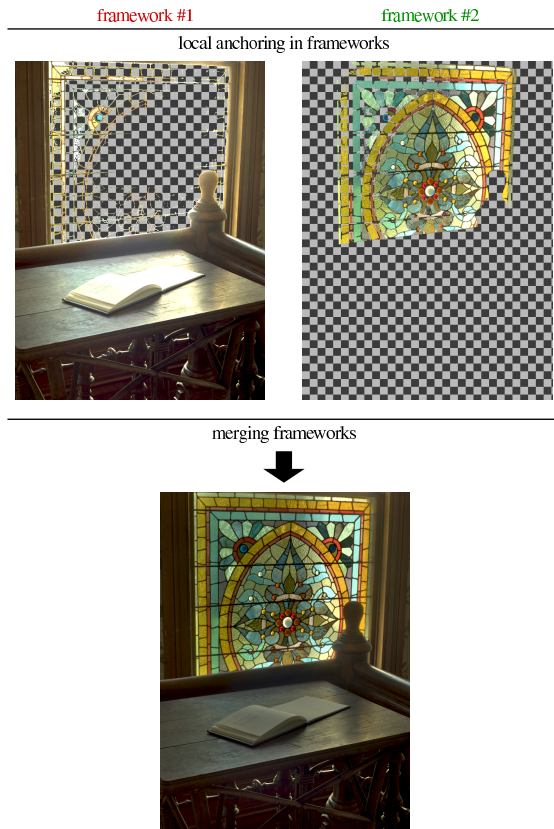


Figure 5: Local anchoring in the frameworks and the final tone mapping result obtained by merging the frameworks.

work i (all these values are in the \log_{10} space), and P_i is the probability map.

In Figure 6 we show how the shifting influences the location of the local anchors and how the merging process affects the image histogram. It is visible in the bottom histogram that all but the darkest areas of the image are fit into the displayable range. The final tone mapped result of the example image (Figure 1) is shown in Figure 5.

5. Results

The main focus of this paper is to demonstrate the application of the lightness perception theory to tone mapping. We first present the results of our method by simulating a well known perceptual experiment. Next, we illustrate the results of the tone mapping of several HDR images with different features and discuss the accuracy of the anchoring and of the frameworks identification. Finally, we comment on the complexity of our algorithm.

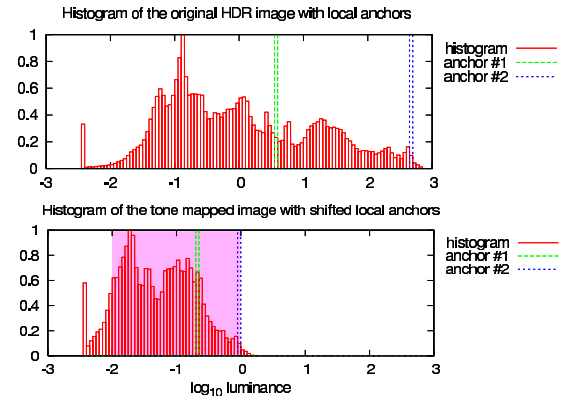


Figure 6: The histogram of the HDR image from Figure 1 with local anchors (top). Below is the histogram of the tone mapped image illustrating how the local anchors are mapped globally. We map the $-2 : 0$ range to displayable values, since a typical display device is capable to display the luminance range of two orders of magnitude (the marked area on the bottom histogram).

5.1. Simulation of Gelb Effect

The Gelb Effect is a well known illusion and provides a good example of lightness constancy failure. This perceptual phenomena is obtained in the following conditions. If in the darkroom with an ambient light a piece of black paper is suspended in the midair and is illuminated by a beam of light, it appears white. However when a piece of real white paper is shown next to the black one, the black paper becomes perceptually gray or black. Apparently the black paper is perceptually darkened by the adjacent paper that has a higher reflectance. Adding subsequent patches of increasing reflectance further darkens the black paper. Clearly, this effect can by definition be attributed to the anchoring in general and to the highest luminance rule in particular. Moreover, this effect cannot be explained with the contrast theories, because the papers do not have to be placed adjacent to each other [GKB⁺99].

We performed a case study of this experiment to validate the results of our algorithm. For comparison, we chose two representative tone mapping algorithms. The *photographic tone reproduction* algorithm presented in [RSSF02], which is based on an image processing technique and follows the anchoring to middle-gray rule, and the *fast bilateral filtering* presented in [DD02], which is related to the intrinsic image models. In the case study, we used five HDR images, each having the same background luminance equal to 10^{-2} , showing from one to five patches with progressively increasing maximum reflectance. The luminance of patches was respectively equal to $[-1, -0.75, -0.5, -0.25, 0]$ in the \log_{10} space.

The results of processing of these images with the chosen tone mapping algorithms are shown in Figure 7. The fast bi-

Tone mapping presented in this paper



Photographic Tone Reproduction [RSSF02]



Fast Bilateral Filtering [DD02]

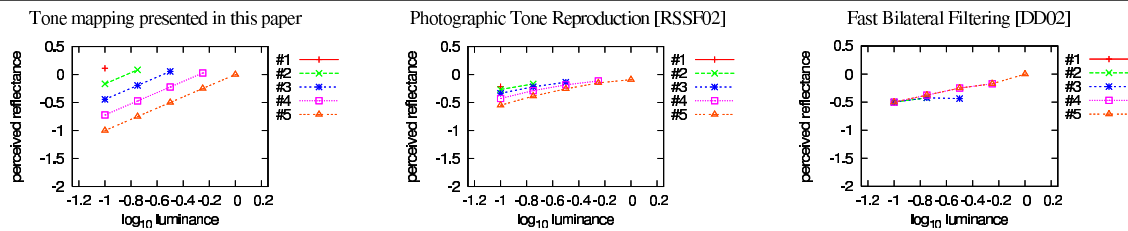


Figure 7: Simulation of the Gelb Effect by various tone mapping algorithms. The plots illustrate how the luminance of the patches is mapped to perceived reflectance in case of each of the five images. On the scale of perceived reflectance, value 0 maps to white, 0.5 to gray and -2 to black. Refer to Section 5.1 for the discussion.

lateral tone mapping [DD02] maps respectively each patch to the same perceived lightness value throughout all five images. This is in accordance with the lightness constancy rule, but contrary to what the observers perceive. Evidently, this method is not able to predict the failure of the lightness constancy rule in this case. Using the photographic tone reproduction [RSSF02] algorithm, the perceived reflectance of the black patch is lowered by the presence of the patches with higher reflectance what follows the Gelb illusion. However in none of the five situations, even with all five patches present, the brightest patch is mapped to a value perceived as white. This inaccuracy happens because the average luminance rule is used for anchoring. Clearly, using our approach we are able to reproduce the same lightness impression as reported for the Gelb experiment.

5.2. Tone Mapping

To illustrate the performance of our tone mapping algorithm, we chose several HDR images, which contain various distinct illumination features. We present our results in Figure 8, where each pane contains a tone mapped HDR image, a map of frameworks areas, and histograms of the original and tone mapped image.

One important issue is that the Gilchrist's model generally assumes approximately diffuse surfaces and if self-luminous areas exist, they occupy a limited field of view. In our application we use this theory for complex scenes be-

yond what was originally tested but we did not observe any problem invalidating our approach. To our knowledge, perceptual model of lightness perception able to deal with natural scenes, which contain large self-luminous surfaces, does not exist.

The decomposition into frameworks obtained using our approach correlates with the intuitive impression of which areas have common illumination. For instance the foreground of the TREE image is in the shadow but the sunny day gives a completely different illumination in the background. In the DESK, CAFE and both OFFICE images the indoor and the outdoor parts are well distinguished even with a grating structure of the window pane in the OFFICE image. Also, the algorithm is able to identify areas illuminated with lights during the night as it can be seen in the FOG and DESK images. In general, the extracted frameworks are plausible despite the lack of semantical information, which might seem to be necessary to perform a successful decomposition. Interestingly, the images tend to be decomposed only into two frameworks, although there is no restriction on their number.

The local and global anchoring of luminance is smoothly modulated by the probability maps of the frameworks. The brightness impression is well reproduced and the anchoring to the highest luminance ensures that the luminance is correctly mapped to lightness. The probability maps have at least a certain minimal influence on each area, thus they pre-

vent the inverse gradients artifacts and the reversal of brightness relation between frameworks.

It is important to note that the spatially processed frameworks do not affect local contrasts during the tone mapping. On the other hand, the linear handling of luminance to merge the frameworks preserves the original colors.

In contrast to the previous approaches to tone mapping that used the segmentation [YP03], we do not aim in our algorithm for estimating the local adaptation level for each pixel. Yee and Pattanaik decompose an image into several layers with different segmentation properties and then average them to finally estimate the local adaptation level for each pixel. On the contrary, in our approach we group the pixels that are under consistent illumination so that they can undergo a common processing.

We avoid the direct picture-to-picture comparison to the previous methods of tone mapping because such a judgement would be highly subjective. However a question arises what improvement our algorithm brings with respect to the sole bilateral filtering [DD02]. The quality of the dynamic range reduction using the bilateral filtering is very high, however the final mapping of the reduced range to the display device is left to the user, whereas our method precisely defines the mapping. Finally, one can benefit from the extraction of the frameworks, which apart from the tone mapping, enable local application of the effects like chromatic adaptation or luminance adaptation.

5.3. Performance

The tone mapping of an HDR image using our algorithm is a matter of seconds on a modern PC. The majority of the computation time is spent on the decomposition into frameworks. Once the frameworks are known, the estimation of anchor and merging of frameworks consists of simple operations. The K-means algorithm operates on a histogram and is therefore independent of the image resolution. The only bottleneck is the spatial processing using the bilateral filter, although we use an efficient approach presented in [DD02].

6. Conclusions

We have presented a novel tone mapping operator which aims at the accurate reproduction of lightness perception of the real world scenes on low dynamic range displays. We leveraged the anchoring theory of lightness perception to handle complex images by developing an automatic method for the image decomposition into frameworks. Through the estimation of the local anchors we formalized the mapping of the luminance values to lightness. The strength of our operator is especially evident for difficult shots of real world scenes, which involve distinct regions with significantly different luminance levels.

As a future work we plan to extend our technique to handle dynamic sequences. The concept of frameworks gives

a unique possibility for the time-dependent local adaptation via the smoothing of the local anchors values. A naive approach to simulate the effect of local adaptation is to smooth the changes of individual pixel values over time, thus simulating the luminance adaptation of the photo receptors. For the moving objects that have a significantly different luminance level than the background, this may lead to the ghosting effect. In fact, the HVS performs a tracking of moving objects of interest with the smooth-pursuit eye movements, therefore the retinal image of these objects is unchanged despite their movement on the display. With the help of frameworks we could follow the objects and perform the local adaptation correctly.

Acknowledgments

We would like to thank OpenEXR, Fredo Durand, Greg Ward, Max Lyons, and Spheron Inc. for making their HDR images available. Special thanks go to Sumant Pattanaik, Rafał Mantiuk, and Philipp Jenke for the valuable discussions and helpful comments concerning this work.

References

- [Are94] AREND L.: *Lightness, Brightness, and Transparency*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994, ch. Intrinsic image models of human color perception, pp. 159–213.
- [Ash02] ASHIKHMEN M.: A tone mapping algorithm for high contrast images. In *Proc. of the 13th Eurographics workshop on Rendering* (2002), pp. 145–156.
- [BT78] BARROW H., TENENBAUM J.: Recovering intrinsic scene characteristics from images. In *Computer Vision Systems*. Academic Press, 1978, pp. 3–26.
- [CM02] COMANICIU D., MEER P.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (2002).
- [DCWP02] DEVLIN K., CHALMERS A., WILKIE A., PURGATHOFER W.: Tone Reproduction and Physically Based Spectral Rendering. In *Eurographics 2002: State of the Art Reports* (2002), Eurographics, pp. 101–123.
- [DD02] DURAND F., DORSEY J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics* 21, 3 (July 2002), 257–266.
- [FLW02] FATTAL R., LISCHINSKI D., WERMAN M.: Gradient domain high dynamic range compression. *ACM Transactions on Graphics* 21, 3 (July 2002), 249–256.

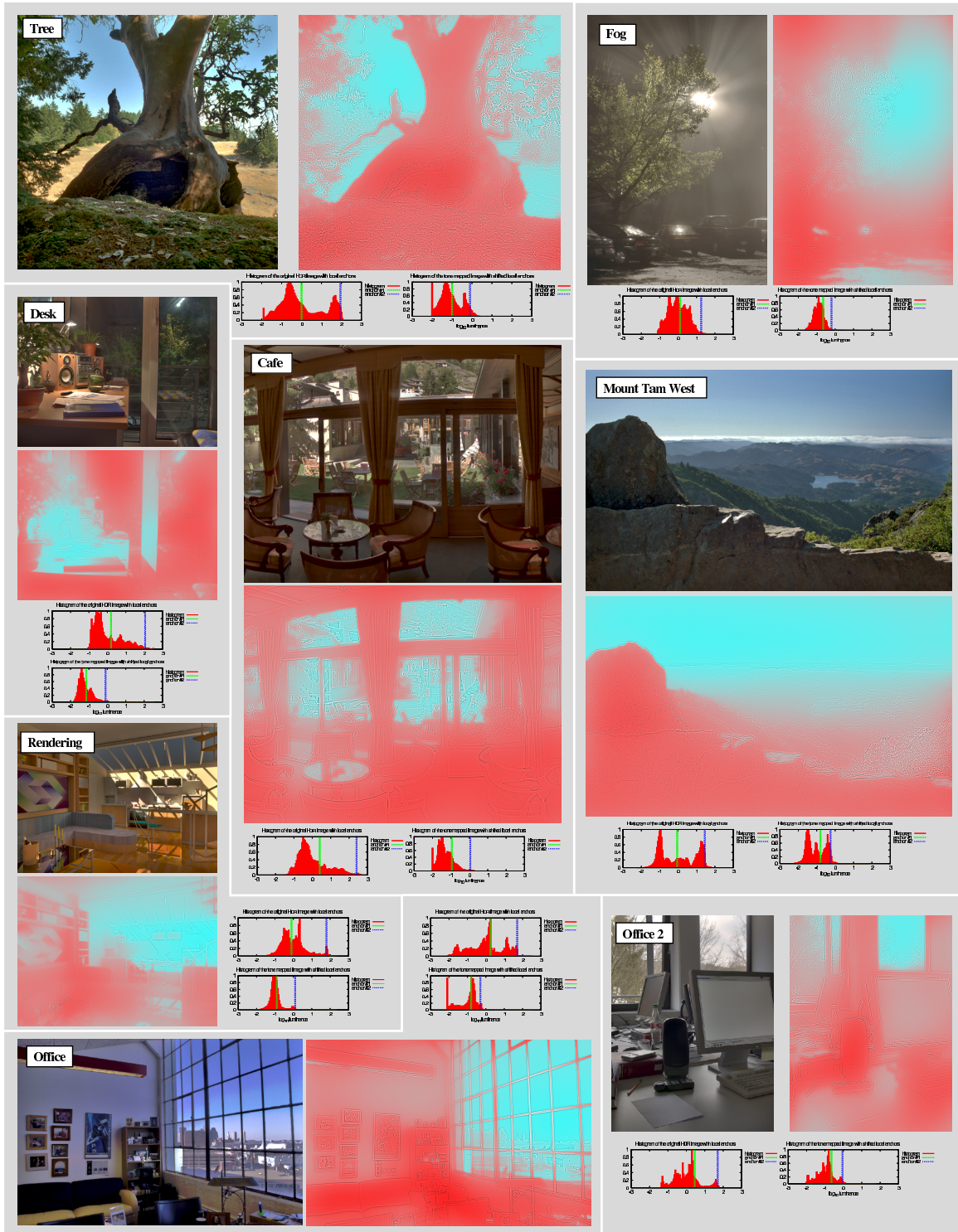


Figure 8: Results of our tone mapping algorithm. Each pane contains a tone mapped HDR image, a map of frameworks area, and histograms of the original (left or top) and tone mapped image (right or bottom). Histograms illustrate the same information as in Figure 6. Two distinguished frameworks are represented by red and cyan colors and are superimposed on an edge filtered image. Higher saturation of the color illustrates stronger anchoring within the framework. Refer to Section 5.2 for the discussion.

- [FPSG96] FERWERDA J., PATTANAİK S., SHIRLEY P., GREENBERG D.: A model of visual adaptation for realistic image synthesis. In *Proceedings of SIGGRAPH 96* (Aug. 1996), Computer Graphics Proceedings, Annual Conference Series, pp. 249–258.
- [GC94] GILCHRIST A., CATALIOTTI J.: Anchoring of surface lightness with multiple illumination levels. *Investigative Ophthalmology and Visual Science* 35 (1994).
- [Gil77] GILCHRIST A. L.: Perceived lightness depends on perceived spatial arrangement. *Science* 195 (1977), 185–187.
- [Gil88] GILCHRIST A.: Lightness contrast and failures of constancy: A common explanation. *Perception & Psychophysics* 43 (1988), 415–424.
- [GKB*99] GILCHRIST A., KOSSYFIDIS C., BONATO F., AGOSTINI T., CATALIOTTI J., LI X., SPEHAR B., ANNAN V., ECONOMOU E.: An anchoring theory of lightness perception. *Psychological Review* 106, 4 (1999), 795–834.
- [Hel64] HELSON H.: *Adaptation-level theory*. New York: Harper & Row, 1964.
- [Hor74] HORN B.: Determining lightness from an image. *Computer Graphics and Image Processing* 3, 1 (1974), 277–299.
- [JRW97] JOBSON D. J., RAHMAN Z., WOODDELL G. A.: A multi-scale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing: Special Issue on Color Processing* 6, 7 (July 1997), 965–976.
- [LG99] LI X., GILCHRIST A.: Relative area and relative luminance combine to anchor surface lightness values. *Perception & Psychophysics* 61 (1999), 771–785.
- [LM71] LAND E. H., MCCANN J. J.: Lightness and the Retinex Theory. *Journal of the Optical Society of America* 61, 1 (1971), 1–11.
- [Pal99] PALMER S.: *Vision Science: Photons to Phenomenology*. The MIT Press, 1999, ch. 3.3 Surface-Based Color Processing.
- [PFFG98] PATTANAİK S. N., FERWERDA J. A., FAIRCHILD M. D., GREENBERG D. P.: A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (1998), ACM Press, pp. 287–298.
- [PTYG00] PATTANAİK S., TUMBLIN J., YEE H., GREENBERG D.: Time-dependent visual adaptation for realistic image display. In *Proceedings of ACM SIGGRAPH 2000* (July 2000), Computer Graphics Proceedings, Annual Conference Series, pp. 47–54.
- [RD05] REINHARD E., DEVLIN K.: Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics* 11, 1 (2005), 13–24.
- [Roc83] ROCK I.: *The logic of perception*. MIT Press, 1983.
- [RSSF02] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: Photographic tone reproduction for digital images. *ACM Transactions on Graphics* 21, 3 (2002), 267–276.
- [SS60] STEVENS S., STEVENS J.: Brightness function: parametric effects of adaptation and contrast. *Journal of the Optical Society of America* 50, 11 (Nov. 1960), 1139A.
- [THG99] TUMBLIN J., HODGINS J. K., GUENTER B. K.: Two methods for display of high contrast images. *ACM Transactions on Graphics* 18, 1 (January 1999), 56–94. ISSN 0730-0301.
- [TM98] TOMASI C., MANDUCHI R.: Bilateral filtering for gray and color images. In *ICCV* (1998), pp. 839–846.
- [TR93] TUMBLIN J., RUSHMEIER H. E.: Tone reproduction for realistic images. *IEEE Computer Graphics and Applications* 13, 6 (Nov. 1993), 42–48.
- [TT99] TUMBLIN J., TURK G.: LCIS: A boundary hierarchy for detail-preserving contrast reduction. In *Siggraph 1999, Computer Graphics Proceedings* (Los Angeles, 1999), Rockwood A., (Ed.), Annual Conference Series, Addison Wesley Longman, pp. 83–90.
- [War94] WARD G.: A contrast-based scalefactor for luminance display. In *Graphics Gems IV*, Heckbert P., (Ed.). Academic Press, Boston, 1994, pp. 415–421.
- [WLRP97] WARD LARSON G., RUSHMEIER H., PIATKO C.: A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics* 3, 4 (1997), 291–306.
- [YP03] YEE Y. H., PATTANAİK S.: Segmentation and adaptive assimilation for detail-preserving display of high-dynamic range images. In *The Visual Computer*, vol. 19(7-8). Springer, 2003, pp. 457–466.