Computational Model of Lightness Perception in High Dynamic Range Imaging

Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel

MPI Informatik, Saarbrücken, Germany

ABSTRACT

An anchoring theory of lightness perception by Gilchrist et al. [1999] explains many characteristics of human visual system such as lightness constancy and its spectacular failures which are important in the perception of images. The principal concept of this theory is the perception of complex scenes in terms of groups of consistent areas (frameworks). Such areas, following the gestalt theorists, are defined by the regions of common illumination. The key aspect of the image perception is the estimation of lightness within each framework through the anchoring to the luminance perceived as white, followed by the computation of the global lightness. In this paper we provide a computational model for automatic decomposition of HDR images into frameworks. We derive a tone mapping operator which predicts lightness perception of the real world scenes and aims at its accurate reproduction on low dynamic range displays. Furthermore, such a decomposition into frameworks opens new grounds for local image analysis in view of human perception.

Keywords: anchoring theory, lightness perception, tone mapping, high dynamic range, segmentation, image processing

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.[1]

The problem of lightness perception and lightness constancy has been studied extensively in the last two centuries (refer to Palmer[1] 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[2] 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 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.[3] This led to the concept of the decomposition of retinal images into the so called intrinsic images[4, 5] 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.[6] 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.[6] which is supported by extensive experimental studies with the human subjects.

Unfortunately, the anchoring theory is not accompanied by any computational model which would enable its direct application. In view of the recent advances in High Dynamic Range (HDR) imaging which allow to capture real world luminance values for further processing,[7,8] it is particularly interesting to develop a computational model of the anchoring theory which would enable its automatic application to natural scenes. The aim of this work is to develop such a model by providing a practical implementation of the key concepts of this theory.

The paper is organized in the following way. We start, in Section 2, with an overview of existing practical implementations of various lightness and brightness perception theories. We then describe the key concepts of the anchoring theory in Section 3 and explain their practical implementation in Section 4. We examine our computational model by reproducing two perceptual experiments in Section 5. Finally we demonstrate several practical applications of our model in Section 6 and conclude our work with future directions in Section 7.

2. PREVIOUS WORK

Lightness and brightness perception issues have been revised while developing the tone mapping operators for HDR images.[9] A contrast ratio present in such images exceeds the capabilities of typical display device, hence it is necessary to map original luminance values to a scale of brightness values that is obtainable on a monitor. Since a luminance match is not possible, it would be desirable to reproduce on a display device a perceptual impression of lightness from the observation of the original scene. This is not an easy task because the reproduction of images on display devices introduces constraints not only in terms of limitations in the luminance range, but also in terms of a narrower field of view and different observation conditions. Such constraints lead to inconsistent behavior of the HVS when viewing real world scene versus their reproductions on a display device.

A number of tone mapping operators have been proposed to overcome these constraints. Initially, the algorithms were based on the power-law relationship between the brightness and the corresponding luminance, as proposed by Stevens and Stevens.[10] 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. Implementations of this approach were presented by Tumblin and Rushmeier[11] (Stevens law) and Drago et al.[12] (Weber-Fechner law). Further attempts in lightness reproduction lead to direct applications of the Retinex theory[2] to tone mapping. Jobson et al.[13] 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,[14] Fattal et al.[15] proposed a successful gradient domain tone mapping operator. A concept of intrinsic images,[4,5] separating the illumination and reflectance (details) layers, inspired many algorithms. The idea was implemented by Tumblin et al.,[16] 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[17] separates the image into large scale features (presumably illumination) and fine details. A much better separation has been achieved using the bilateral filter.[18]

The early methods based on a uniform mapping using brightness functions suffered from the loss of fine details while reproducing brightness to a limited contrast range. This was improved by contrast techniques based on Retinex theory, however these contain an integration step which produces a result image with an unknown offset. Thus, it is not clear how to map the result on the scale of perceived gray levels. On the other hand, the intrinsic image models fail to address the apparent failures in lightness constancy. These arguments served as a motivation to develop a computational model of currently the most advanced theory of lightness perception. The first attempt to use the basic concepts of this theory has been presented by Krawczyk et al.[19] in view of tone mapping application. Here we improve the computational model by including a more complete set of concepts and examine it with a perceptual experiment in the context of natural scenes.

3. ANCHORING THEORY OF LIGHTNESS PERCEPTION

The anchoring theory of lightness perception by Gilchrist et al.[6] is qualitatively different from the recent lightness models and is based on a combination of global and local anchoring of lightness values. In the following sections we explain the main concepts of this theory. First, we discuss the estimation of lightness within the simple scenes (background/patch stimuli) using the anchoring rules. Next, we explain how to extend the anchoring of lightness to the complex scenes using frameworks.

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 mapping is referred to as *scaling*. Although usually a veridical scaling is assumed, the compression or expansion of range is possible if necessary.

The anchor cannot be defined once for the absolute luminance values, because each luminance level can be perceived as any shade of gray depending on the observation conditions. It therefore must be tied to some measure of relative luminance values. According to Li et al.,[20] such a measure is properly defined by the highest luminance rule affected by the relative area. 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 evident perception of self-luminous surfaces (lighter than white) lead to an extended definition. Apparently, there is a tendency of the highest luminance to appear white and a tendency of the largest area to appear white. When the highest luminance covers the largest area, the highest luminance becomes a stable anchor and is mapped to white. However, if the darker area becomes larger, the highest luminance starts to be perceived as self-luminous and the anchor becomes a weighted average of the luminance proportionally to the occupying area.

An alternative measure is defined by the average luminance rule[21] and states that the average luminance in the visual field is perceived as middle gray and serves as a standard for mapping of both lighter and darker surfaces. The experimental evaluation of the average luminance rule versus the highest luminance rule decisively favors the latter.[20] Other findings[22] based on more complex stimuli agree with these conclusions.

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.[6] introduce the concept of decomposition of an image into components, *frameworks*, in which the anchoring rule can be applied directly. In their 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. Additionally, proximity is also considered as a grouping factor. 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. A target in a complex image, that belongs to more than one framework, may have different lightness values when anchored within different frameworks. The influence on the final lightness of a target is decided depending on the strength of each framework. The strength of a framework is mainly determined by its size and the variety of luminance values it contains – frameworks with lower variance or smaller in size are weaker.

Besides the local frameworks, the whole image constitutes an additional *global framework* with its global anchor. 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 local frameworks in proportion to the strength of these frameworks and with certain constant level of influence of the global framework.

4. COMPUTATIONAL MODEL

The anchoring theory has been presented without a formal model. On a technical level this requires to develop a method for automatic decomposition of an image into frameworks, to build heuristics estimating the influence of each framework on total lightness, and to estimate the anchors within the frameworks. Furthermore, the algorithm must perform accurately when used with natural scenes.

The presented model takes an image with the relative luminance values as an input. Such values can be computed from RGB channels of an HDR image according to CIE XYZ luminous efficiency functions. We first decompose the input image into the overlapping frameworks and define the probabilities with which pixels belong to each framework. Next, we estimate the anchor in each framework, i.e. the luminance value perceived as white. The frameworks and anchors allow us to calculate the net lightness of the scene.

4.1. Decomposition into Frameworks

We define a framework as a probability map over the whole image in which a probability of belonging to a framework is assigned to each pixel. A framework, in order to be *valid*, must be defined by a number of pixels which belong to this framework with a probability above 95%. In principle, the task of decomposition into frameworks is to identify luminance values which represent the potential frameworks in most accurate way and to assign appropriate probability values for each pixel.

It is reasonable to think of the decomposition as a segmentation problem. We experimented with several segmentation algorithms in order to find a plausible decomposition into frameworks, and the mean shift segmentation[23] produced the most appropriate results. However segmentation algorithms in general assign a pixel to only one segment and therefore do not implement the notion of probability in belonging to a segment. A border between two frameworks which might occur on a smooth gradient in the image is in such situation impossible to represent correctly. We therefore decided to tailor a custom decomposition method.

Our method is based on the luminance intensities in the input image, which is sufficient since we define frameworks in terms of areas of common illumination. 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 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. Upon convergence, we remove centroids representing empty segments.

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 *belongingness* to the centroid with a Gaussian function:

$$P_i(x,y) = e^{\frac{-(C_i - Y(x,y))^2}{2\sigma^2}} \tag{1}$$

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.

Often at this stage, several centroids may represent an invalid framework (none of the pixels belongs to it with probability above 95%). We iteratively merge these centroids. In each iteration, the centroid representing an invalid framework and the adjacent closest centroid are merged together and the new centroid value is equal to their weighted average proportional to their area:

$$C_{i,j} = \frac{C_i \cdot S_i + C_j \cdot S_j}{S_i + S_j} \tag{2}$$

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 these centroids. After the merge, probability values are recalculated according to (1), and the iteration is repeated until no centroids need to be merged.

As the next step, we spatially process the probability map of each framework to include the *proximity* aspect of Gestalt grouping factors. The spatial processing smoothes local variations in the probability values which may appear due to a texture on surfaces. However, it is important to preserve high local variations which could appear on the borders of objects or frameworks. The bilateral filter[24] 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.

We demonstrate the decomposition procedure on an example HDR image shown in Figure 1(a). First, we converge the initial segmentation (Figure 2(a)) to identify luminance values that would represent the most accurate decomposition into frameworks. We then calculate the probabilities (Figure 2(b)) and merge one centroid which does not represent a valid framework. The final centroids (Figure 2(c)) define the probability maps based only on luminance property (Figure 1(b)) and contain several incorrect assignments. The reflections on the logs at the bottom of the image are incorrectly assigned to a framework which mainly contains clouds. We refine the probability maps to include also spatial interactions as shown in Figure 1(c).
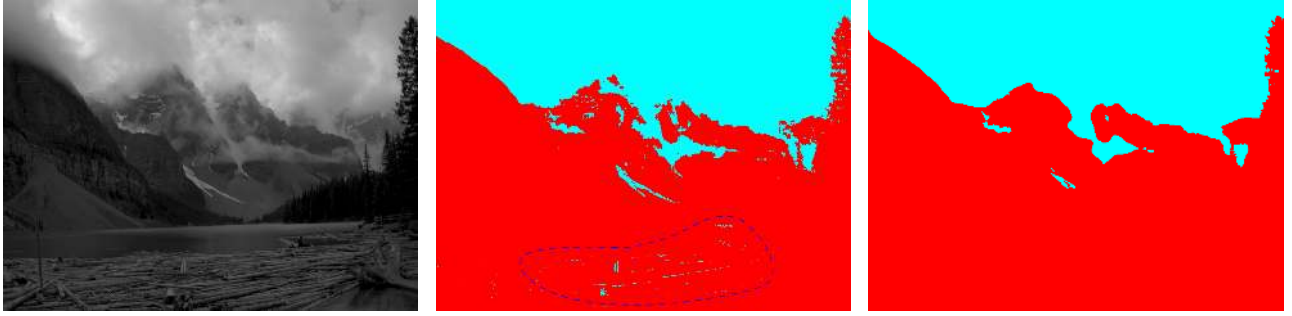
(a) luminance channel of example HDR image

(b) intensity based frameworks

(c) spatially processed frameworks

Figure 1. MORAINE – an example HDR image courtesy of Greg Ward. The image appears dark, because it has been exposed for details in the clouds. Figures 1(b) and 1(c) illustrate the decomposed frameworks before and after spatial processing. Notice the artifacts in the intensity based frameworks decomposition (marked area) which are corrected after the spatial processing. Refer to Section 4.1.
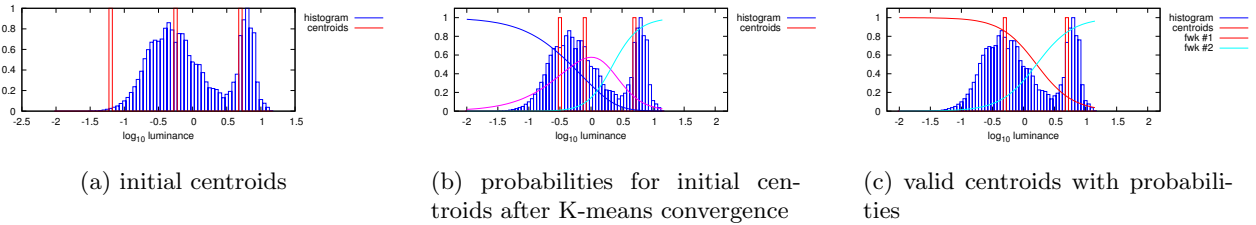
(a) initial centroids

(b) probabilities for initial centroids after K-means convergence

(c) valid centroids with probabilities

Figure 2. The histogram of the HDR image from Figure 1(a) illustrating the estimation of centroids which provide an appropriate decomposition into frameworks. In the middle and right 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. Refer to Section 4.1.

4.2. Strength of Framework

Apart from the illumination conditions, the strength with which a framework influences the lightness of a given surface also depends on the articulation of the framework and its relative size. If a framework is highly articulated, the pixels tend to be strongly anchored within this framework. Also, large frameworks have higher influence on the lightness than small ones. A strength of the framework is defined as a product of the articulation and size factors.

We estimate the articulation factor 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:

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

(3)

where A_i denotes the articulation factor 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.

Similarly, a larger framework will have a tendency to a higher influence on the lightness of surfaces, while a relatively small framework will have a rather limited impact. We estimate the size factor in the following way:

$$X_i = 1 - e^{\frac{-(S_i)^2}{2 \cdot 0.1^2}}$$

(4)

(a) local anchoring for framew. #1   (b) local anchoring for framew. #2   (c) net lightness
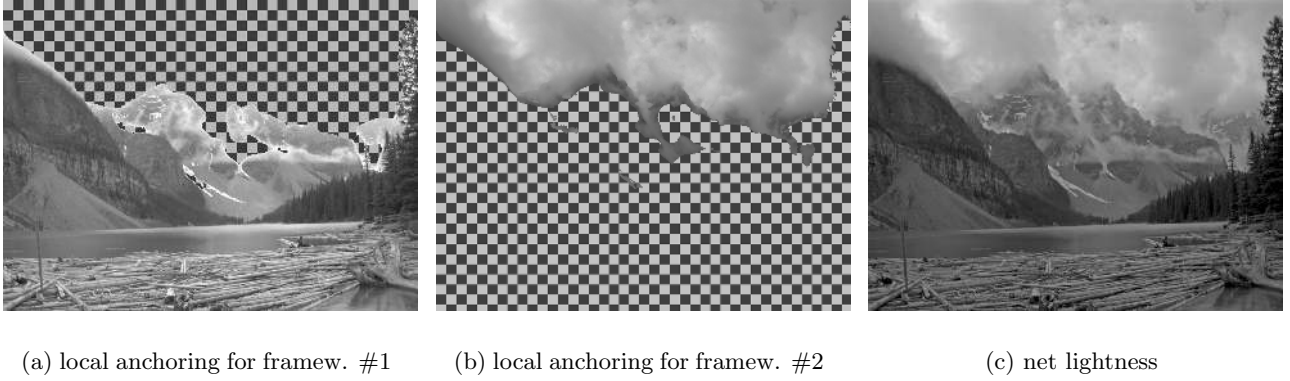
Figure 3. Local anchoring in the frameworks and the net lightness calculated as described in Section 4.4. Compare the result in Figure 3(c) to the input in Figure 1(a).

where X_i denotes the size factor of the framework i, and S_i represents the normalized relative area of a framework. Here we attenuate the influence of frameworks whose size is below 30% of the total image area.

We apply the strength factor to frameworks by multiplying their probability maps P_i by their respective articulation factor A_i and size factor S_i. We then normalize the probability maps again and obtain the final result of the decomposition into frameworks.

The parameters used to estimate the influence of size and articulation on the strength of the framework are chosen arbitrarily. However, it is not as important to have a precise estimation as to include such factors at all. The suggested values served well in our experiments and fine tuning leads to rather marginal changes. Yet, failure to use these factors leads to an excessive influence of unimportant frameworks on the net lightness estimation.

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.

Although we apply the highest luminance rule, we cannot directly use the highest luminance in the framework as an anchor. As discussed in Section 3.1, there is a relation between what is locally perceived as white and its area relatively to the surround. This implies that a spatial filtering is required prior to the estimation of the local anchor. Our procedure is to filter the area of a framework with a large Gaussian kernel to suppress plausible small areas with high luminance. We thus eliminate potential self-luminous areas, what allows us to take the highest luminance of the rest of the pixels as the anchor. In Figures 3(a) and 3(b) we show the two frameworks identified in the example HDR image with their lightness computed according to the local anchor.

4.4. Net Lightness

Given the decomposed frameworks and estimated local anchors we compute the net lightness of the pixels by merging the frameworks. We process each framework individually. We sum the original luminance values of the HDR image normalized by the locally estimated anchor value and proportionally to the probability map:

$$L(x,y) = 30\% \cdot \sum_i (Y - W_i) \cdot P_i(x,y) + 70\% \cdot (Y - W_0) \tag{5}$$

where L denotes the final lightness value, Y the original luminance of the HDR image, W_i the local anchor of framework i, W_0 the anchor in the global framework (all these values are in the $\log_{10}$ space), and P_i is the probability map. The 30% and 70% coefficients for local and global anchor influence respectively are arbitrarily suggested by Gilchrist et al.[6] and can be modified if necessary.

| (a) original photo (cropped) | (b) identified frameworks | (c) probe disks of const. luminance |

Figure 4. The decomposition into frameworks within a multi-illuminant natural scene and a reproduction of Gilchrist and Radonjic experiment.[25] Refer to Section 5.1. Original photo by Henri Cartier-Bresson.

In Figure 3(c) we illustrate how the net lightness has been computed for sample HDR image. Comparison to the original HDR image from Figure 1(a) clearly illustrates an improved perception of surface details in the processed image.

4.5. Performance of The Model

The estimation of lightness in a 4Mpx image using our computational model takes below a minute on a modern PC. The timing mainly depends on the number of decomposed frameworks, since the majority of computations is spent on the decomposition stage. Once the frameworks are known, the estimation of anchor and net lightness computation consist 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 by Durand and Dorsey.[18]

5. TESTING THE MODEL

The main focus of this paper is a computational model of the lightness perception theory. A thorough verification of the presented model would require a psychophysical experiment which is beyond the scope of this paper. Instead we test our computational model by simulating two perceptual experiments. The first one analyzes the accuracy of decomposition into frameworks for natural scenes and the second experiment is a simulation of the Gelb illusion using various lightness mapping algorithms for HDR images (tone mapping operators).

5.1. Frameworks within Multi-Illuminant Scenes

According to the anchoring theory of lightness perception, successfully identified frameworks should define the areas in which the lightness is perceived homogeneously.[6] An evidence for such lightness perception can be obtained through a distribution of probe disks of constant known luminance value across the image. The disks should have the same lightness within a framework independently of the ratio of their luminance to the background luminance on which they are placed.

Such an experiment has recently been presented by Gilchrist and Radonjic.[25] Here we provide an HDR reproduction of this experiment (Figure 4) using the original material. We revert the original photo to a high dynamic range image through an assumption on the camera response.[8] We then decompose the HDR image into frameworks using presented computational model (Figure 4(b)) and place several probe disks of constant luminance value in various areas of the image (Figure 4(c)). The contrast ratios between the probe disks and the background in both shadow and light frameworks range from 1:2 to 1:9.
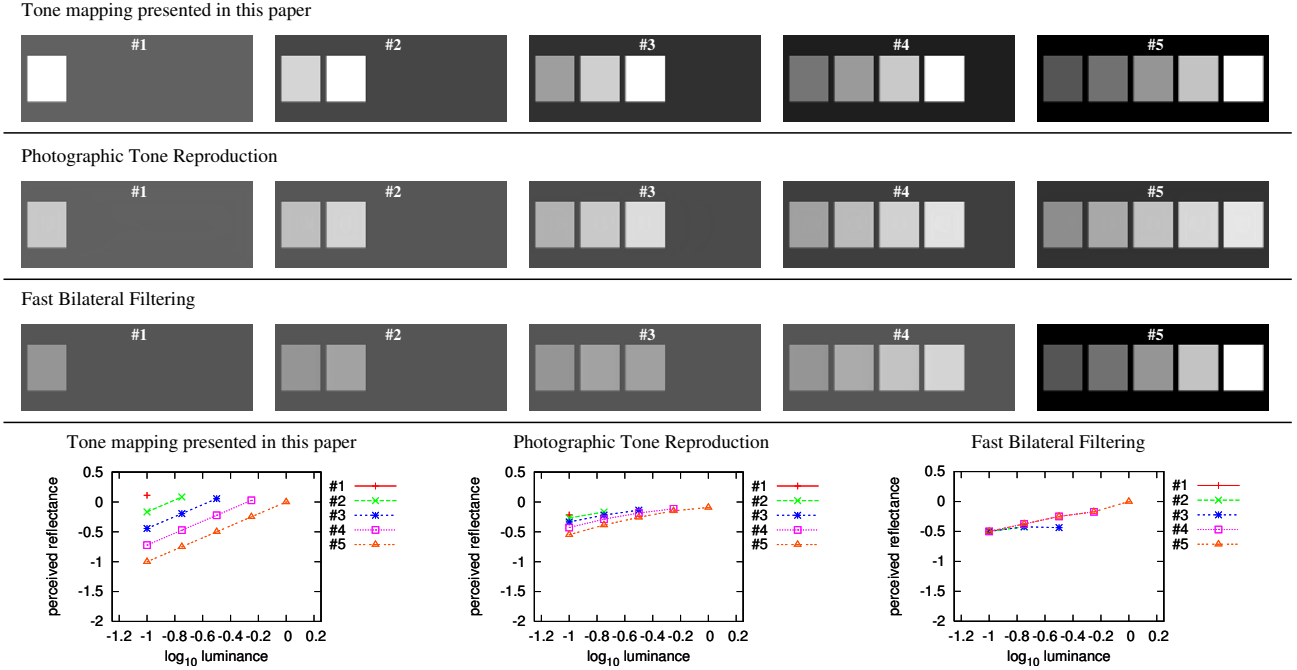
Figure 5. Simulation of the Gelb Effect by various tone mapping algorithms. The plots illustrate how the luminance of the patches is mapped to perceived reflectance in case of each of the five images. On the scale of perceived reflectance, value 0 maps to white, 0.5 to gray and -2 to black. Refer to Section 5.2 for the discussion.

The lightness of the probes is perceived consistently within the area of the frameworks independently of the background. These leads to the same conclusions as in the original experiment: lightness is determined by the frameworks and the influence of local contrasts is minimal. Our contribution here is not to confirm the theory, but to provide an automated method for providing an appropriate decomposition. In this sense, the reproduction of the experiment provides the evidence that frameworks areas are accurately identified using our computational model.

5.2. Anchoring in Gelb Illusion

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.[6]

We performed a case study of this experiment to validate the results of our algorithm. For comparison, we chose two other methods whose principle goal includes correct representation of lightness in the original image. These are tone mapping algorithms for HDR images: the *photographic tone reproduction* algorithm presented by Reinhard et al.[26] is based on an image processing technique and follows the anchoring to middle-gray rule, and the *fast bilateral filtering* presented by Durand and Dorsey,[18] 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 5. The fast bilateral tone mapping[18] 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[26] 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. However, using our computational model it is possible to reproduce the same lightness impression as reported for the Gelb experiment.

6. APPLICATIONS

The computational model of lightness perception theory opens new possibilities in processing of HDR images. As a direct application we present a tone mapping operator which aims at the reproduction of lightness as closely as possible to the lightness perceived in a natural scene. Besides, decomposition into frameworks defined by homogeneous illumination gives an interesting possibility to perceptually supported image processing.

6.1. Tone Mapping

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 based on the luminance channel and builds upon the computational model described in Section 4.

The particular goal of tone mapping is the contrast reduction in the original HDR image. While through the net lightness computation the dynamic range in the image is reduced, it may still exceed the capabilites of the display device. Hence, an additional dynamic range reduction may be necessary to achieve good results. For the purpose of tone mapping we use the modified version of net lightness computation (5) which includes a dynamic range reduction factor D_i:

$$L(x,y) = 70\% \cdot \sum_i \frac{(Y - W_i)}{D_i} \cdot P_i(x,y) + 30\% \cdot (Y - W_0).$$ (6)

The value of D_i is chosen individually for each framework in such way that it scales down the dynamic range of a framework if it exceeds the capabilites of the target display device. Also the influence of the global framework has been limited, because it counteracts the luminance range compression goal.

We first calculate the luminance from the RGB colors using the CIE Yxy color space and decompose the input scene into the frameworks. 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 using equation (6). We recover the color information with an inverse CIE Yxy transform, using the computed lightness L instead of luminance channel Y. The result is suitable to be viewed on an LDR display device.

The evaluation of aesthetic properties of this tone mapping can be done with recently presented methodology.[27, 28] In Figure 6, we provide analysis of how the luminance values are mapped to the lightness levels in three different tone mapping techniques: the presented algorithm, and the global version of *Photographic Tone Reproduction*[26] which is similar to brightness mapping functions, and the *Fast Bilateral Filtering*[18] which is inspired by the intrinsic image models. The key aspect of any tone mapping algorithm is the efficient use of the available limited dynamic range on a display device. Usually the contrast compression results in the loss of fine details when uniform mapping is used (Figure 6(a)). The decomposition of an HDR image into the illumination and reflectance layers leads to a better detail preservation as visible in Figure 6(b). The mapping function in this case is not uniform and maps source luminance to different lightness values depending on the surround, hence makes more efficient use of the available dynamic range. However, the independent estimation of lightness within each framework leads to the most efficient mapping in this case, without invalidating the consistency of
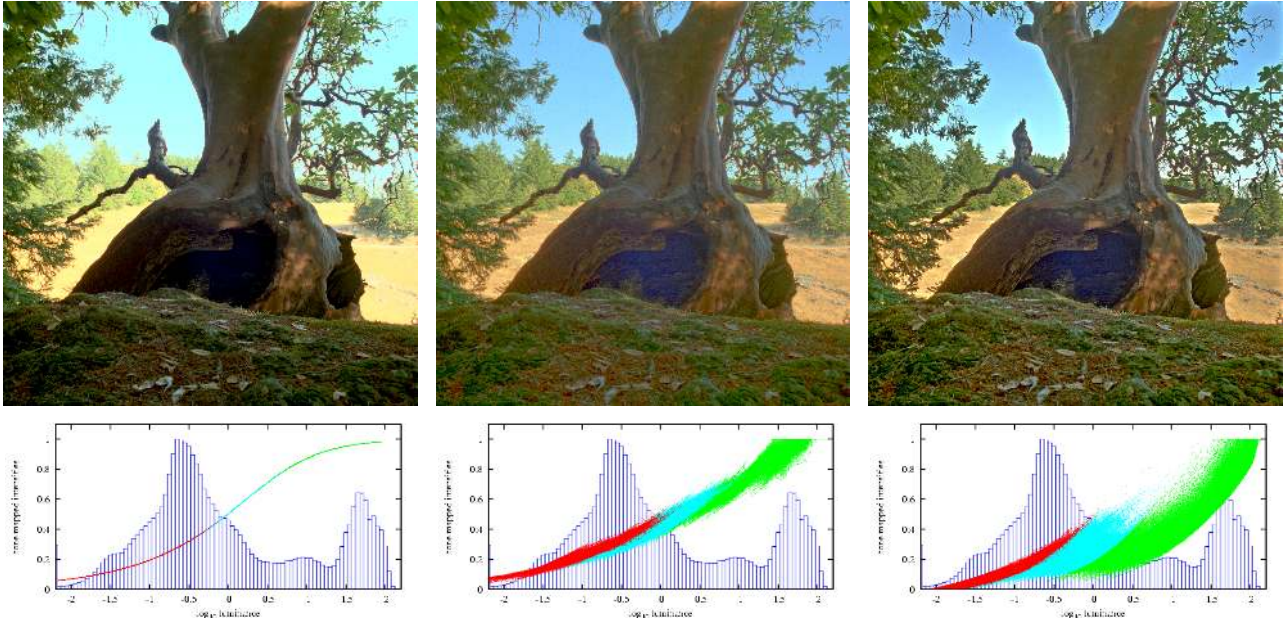
(a) global version of Photographic Tone Reproduction

(b) Fast Bilateral Filtering

(c) presented method
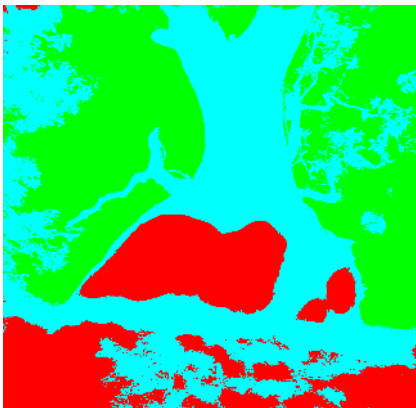
(d) frameworks map

(e) presented method with global white balance

(f) presented method with local white balance in frameworks

Figure 6. TREE scene. Comparison of various tone mapping operators, color of plots of mapping functions represent areas of frameworks. Refer to Section 6.1. Figures 6(e) and 6(f) illustrate global versus local white balance correction, refer to Section 6.2. HDR image courtesy of OpenEXR.

illumination in the scene (Figure 6(c)). Also, the representation of frameworks as probability maps ensures that no artifacts are introduced on the borders of frameworks.

6.2. Local Image Processing

The image processing algorithms are usually applied with uniform parameter settings over the whole image. Even if the algorithm is localized, the parameters for the method are set globally. However, it is often desired

to vary the parameters of such algorithms between the areas of an image. In this case it often has to be done manually by the user. Except that with the use of frameworks decomposition, it is now possible to identify the areas that are perceived homogeneously in an image. For the purpose of automated image processing, this permits to estimate the most appropriate parameters for given algorithm individually for each framework.

In digital photography, it often happens that an image contains two different sources of illumination – for instance a warm daylight and bluish shadow as in Figure 6(c). Such an image requires a white balance correction. However, correcting for the daylight will result in an increased bluish cast in the shadow part as shown in Figure 6(e). The decomposition into frameworks allows to identify multiple illumination types and enables different white balance correction in each of these areas, Figure 6(f). Again, frameworks represented as probability maps guarantee proper blending of edges where differently processed areas merge.

7. CONCLUSIONS

We have presented a computational model of an anchoring theory of lightness perception. The model provides a practical implementation of the key concepts of this theory and aims at an accurate estimation of lightness in real world scenes captured as HDR images. We leveraged the theory 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. We examined the accuracy of our model by reproducing two perceptual experiments that were initially conducted to prove the accuracy of the theory.

We demonstrated a direct application of our computational model to the tone mapping of HDR images and discussed its advantages. Moreover, the decomposition of an image into the frameworks gives additional potential to automated image processing fine tuned to the perceptual aspects of HVS.

ACKNOWLEDGMENTS

REFERENCES

1. S. Palmer, *Vision Science: Photons to Phenomenology*, ch. 3.3 Surface-Based Color Processing. The MIT Press, 1999.

2. E. H. Land and J. J. McCann, "Lightness and the Retinex Theory," *Journal of the Optical Society of America* **61**(1), pp. 1–11, 1971.

3. A. L. Gilchrist, "Perceived lightness depends on perceived spatial arrangement," *Science* **195**, pp. 185–187, 1977.

4. H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*, pp. 3–26, Academic Press, 1978.

5. L. Arend, *Lightness, Brightness, and Transparency*, ch. Intrinsic image models of human color perception, pp. 159–213. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994.

6. A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J. Cataliotti, X. Li, B. Spehar, V. Annan, and E. Economou, "An anchoring theory of lightness perception," *Psychological Review* **106**(4), pp. 795–834, 1999.

7. P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of SIGGRAPH 97, Computer Graphics Proceedings, Annual Conference Series*, pp. 369–378, Aug. 1997.

8. G. Krawczyk, M. Goesele, and H.-P. Seidel, "Photometric calibration of high dynamic range cameras," Research Report MPI-I-2005-4-005, Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany, April 2005.

9. E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, Morgan Kauffman, 2005.

10. S. Stevens and J. Stevens, "Brightness function: parametric effects of adaptation and contrast," *Journal of the Optical Society of America* **50**, p. 1139A, Nov. 1960.

11. J. Tumblin and H. E. Rushmeier, "Tone reproduction for realistic images," *IEEE Computer Graphics and Applications* **13**, pp. 42–48, Nov. 1993.

12. F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Computer Graphics Forum, proceedings of Eurographics 2003* **22**(3), pp. 419–426, 2003.

13. D. J. Jobson, Z. Rahman, and G. A. Woodell, "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), pp. 965–976, July 1997.

14. B. Horn, "Determining lightness from an image," *Computer Graphics and Image Processing* **3**(1), pp. 277–299, 1974.

15. R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Transactions on Graphics* **21**, pp. 249–256, July 2002.

16. J. Tumblin, J. K. Hodgins, and B. K. Guenter, "Two methods for display of high contrast images," *ACM Transactions on Graphics* **18**, pp. 56–94, January 1999. ISSN 0730-0301.

17. J. Tumblin and G. Turk, "LCIS: A boundary hierarchy for detail-preserving contrast reduction," in *Siggraph 1999, Computer Graphics Proceedings*, A. Rockwood, ed., *Annual Conference Series*, pp. 83–90, Addison Wesley Longman, (Los Angeles), 1999.

18. F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Transactions on Graphics* **21**, pp. 257–266, July 2002.

19. G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Lightness perception in tone reproduction for high dynamic range images," in *The European Association for Computer Graphics 26th Annual Conference EUROGRAPHICS 2005, Computer Graphics Forum* **24**(3), pp. xx–xx, Blackwell, (Dublin, Ireland), 2005.

20. X. Li and A. Gilchrist, "Relative area and relative luminance combine to anchor surface lightness values," *Perception & Psychophysics* **61**, pp. 771–785, 1999.

21. H. Helson, *Adaptation-level theory*, New York: Harper & Row, 1964.

22. A. Gilchrist and J. Cataliotti, "Anchoring of surface lightness with multpile illumination levels," *Investigative Ophthamalmology and Visual Science* **35**, 1994.

23. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(5), 2002.

24. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV*, pp. 839–846, 1998.

25. A. Gilchrist and A. Radonjic, "Probe disks reveal framework effects within multi-illuminant scenes," 2005. Eurpoean Conference on Visual Perception, A Coruña, Spain.

26. E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics* **21**(3), pp. 267–276, 2002.

27. P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen, "Evaluation of tone mapping operators using a high dynamic range display," in *ACM SIGGRAPH 2005, LA.*, ACM Press, August 2005.

28. A. Yoshida, V. Blanz, K. Myszkowski, and H. s.-P. Seidel, "Perceptual evaluation of tone mapping operators with real-world sceness," in *Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symp osium on Electronic Imaging (2005)*, pp. 192–203, SPIE, (San Jose, USA), 2005.