Investigation of image perception under low light conditions

Bryce Yao
Stanford University

Tien-Ning Hsu
Stanford University

Abstract

In this paper, we take a closer look at the visual perception difference when the same content is presented at different luminance levels. Different visual perception and color appearance models are discussed and applied to guide the enhancement of an image when it is viewed at a lower luminance level. Two algorithms are further investigated, implemented, and compared using their result images.

1. Introduction

Modern displays are used in different lighting conditions with the awareness of the ambient illumination levels, i.e., the displays automatically adjust the brightness levels depending on the ambient illumination. Such adjustment can reduce energy consumption and produce better visual experience for the users. However, there is a discrepancy that the display content is the same across the range of ambient illuminations, yet human visual systems do not retain a constant color and contrast perception. Therefore, the image quality and visual experience differ drastically when the same picture is shown at different situations. It is the goal of this project to take a closer look at different algorithms to adjust the image based on the ambient light situations and propose the best solution.

2. Related Work

The CIE Technical Committee proposed a color appearance model CIECAM02 [1], based primarily on a set corresponding colors experiments and a collection of color appearance experiments. This model proposed a method to take into account the viewing condition to transform tristimulus values to or from perceptual attribute correlates. Shih et al [2] utilized this color appearance model to preserve image hue, chroma, and brightness under low light conditions. However, Wanat et al [3] argue that CIECAM02 model is almost entirely based on the cone-mediated vision, while a big portion of the color gamut in modern displays often lies in the luminance range below 3 cd/m^2, when visual signal is mainly influenced and perceived by the retinal rod cells. Thus, they propose a new appearance matching model and luminance retargeting method. This method relies on a model of color as well as human contrast sensitivity function [4]. The best compromise between retaining contrast and brightness is provided by optimizing the shape of a tone curve.

3. Method and materials

In this work, two promising methods of compensating images when viewed under low light situations are investigated and compared. Each method is first explained in this section and then implemented in MatLab. Three srgb encoded images are then processed by each of the two algorithms. The result images are presented and discussed in section 4.

Contrast and color retargeting

This method treats contrast and color separately, based on the contrast and color responses under different luminance levels. The flow chart of this algorithm is shown in figure 1. The image is first transformed into YCbCr space to isolate the intensity channel. Both global and local contrast are modified in this layer. When adjusting the local contrast, the image is decomposed into different frequency bands, and recomposed together with different weights calculated based on the spatial frequency and its corresponding contrast threshold. The color saturation and hue are also modified based on the saturation matching functions, human eye absorption functions, and display spectrum power density functions.

3.1.1 Contrast sensitivity model

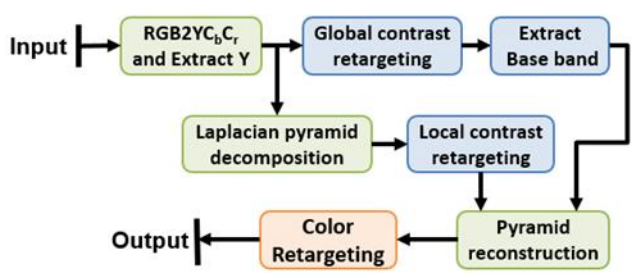

Figure 1: flow chart of Contrast and color retargeting method from [3]

In order to create a consistent visual perception, it is important to understand the visual response functions under different luminance. It has been demonstrated that luminance affects both color and contrast sensitivity of the eye [5]. Many studies [6, 7, 8] have shown that contrast sensitivity functions (CSF) can depend on many factors such as age of the viewer, viewing distance, source of the stimulus, etc. However, for the same image being viewed at two different lighting conditions, it is mainly affected by the luminance levels. A contrast sensitivity function model is studied and recreated (figure 2) from [5]. This CSF function serves as a reference in matching two contrast levels in this algorithm.

3.1.2 *Global contrast retargeting*

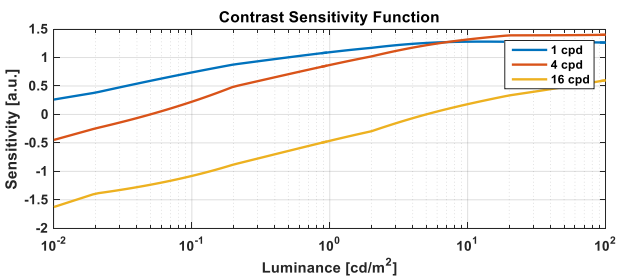

Figure 2: Contrast sensitivity function (CSF) as a function of luminance.

The global contrast is adjusted by a tone curve. It reshapes the general image appearance. This is achieved by a luminance dependent function T(Y). T(Y) maps an input luminance Y to an output luminance $\tilde{Y}$. This tone curves follows the principle that two contrast magnitudes match in appearance if the differences to their respective contrast threshold is the same:

$$M - M_t = \widetilde{M} - \widetilde{M_t} \quad \left(M_t = \frac{1}{S \cdot CSF(\rho, L_a)}\right)$$

Where M and M_t denote a contrast magnitude and the contrast sensitivity with spatial frequency ρ and adaptation luminance L_a. The terms with hats denote the same quantities at target luminance. This T(Y) curve is found by solving the optimization problem

$$\arg\min_{T(l)} \int_{l_{min}}^{l_{max}} S(l)\left(G - G_t(l) - \frac{dT}{dl}G + G_t(T(l))\right)^2 + \\ + \tau\,(l - T(l))^2\, dl$$

$$\text{subject to:} \quad \frac{dT}{dl} \geq 0, \\ T(l_{min}) \geq d_{min}, \quad T(l_{max}) \leq d_{max}.$$

Where l is the luminance Y in log space:

$$l = \log_{10} Y$$

A tone curve to retarget from 100 cd/m² down to 1cd/m² is solved and calculated in figure 3.

3.1.3 *Local contrast retargeting*

Although the appearance of the retargeted image can be

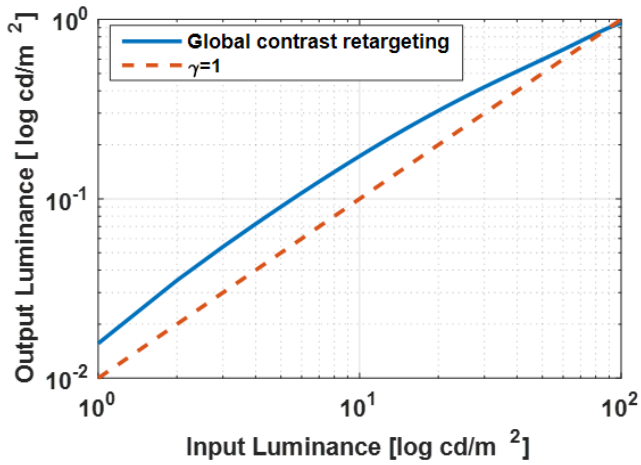

Figure 3: Tone curve for global contrast retargeting (blue solid curve). The luminance level is retargeted from 100 cd/m² to 1 cd/m². A linear scaling curve with gamma =1 (orange dashed curve) is also shown as a reference.

hugely improved by a well-selected tone curve, the contrast matching model still significantly depends on two other parameters: spatial frequency and contrast magnitude. These two properties must be addressed on a local level.

More specifically, the log of the luminance values are extracted from the input image, and this luminance layer is then decomposed into frequency-selective bands (P_k) by Laplacian pyramid decomposition. To construct the Laplacian pyramid, three gaussian kernels with different standard deviation are applied to the image, and the difference between each blurred image becomes the Laplacian pyramid layer. One thing to notice is that the gaussian window should be smaller for higher frequencies to get finer details. The standard deviation of the gaussian kernel is defined as below, and k=1 denotes the finest level.

$$\sigma = 2^k, k = 1, 2, 3$$

After decomposition by Laplacian pyramid, the retargeting image is constructed with different weights computed from the local contrast (m_k):

$$\widetilde{P_k}(x, y) = P_k \cdot m_k(x, y)$$

$$m_k(x, y) = \frac{\sqrt{(g_\sigma * [l(x, y) - (g_\sigma * l)(x, y)]^2)(x, y)} - G(M_t) + G(\widetilde{M_t})}{\sqrt{(g_\sigma * [l(x, y) - (g_\sigma * l)(x, y)]^2)(x, y)}}$$

where g_σ is a Gaussian kernel with standard deviation σ. M_t and $\widetilde{M_t}$ are the detection thresholds for the input and retargeted images. Since the luminance of retarget image is still unknown, the latter is provided by the base-band image from the result of global contrast.

Figure 4: Retargeting results comparison, from top to bottom: original, enhanced after global contrast retargeting, after local contrast retargeting, and after color retargeting.

Finally, these modified layers of the pyramid ($\widetilde{P_k}(x,y)$) are summed together including the base-band image from global contrast. $\tilde{Y}$ can be extracted from the resulting image.

3.1.4 Local contrast retargeting

When luminance is low, the image loses color saturation, along with a hue shift. This is mostly caused by the varying contribution ratio between cones and rods. First, the hue shift is addressed by converting the linear RGB values into photoreceptor responses with matrix M_E, where the coefficients of the matrix are given by the spectral primaries for the display (π_i) multiplied by the spectral sensitivity of each type of photoreceptor (σ_p) summing across all wavelengths. That is,

$$m_{P,i} = \int_{\lambda} \pi_i(\lambda)\, \sigma_p(\lambda) d\lambda$$

Cao et al. [9] observed that the influence of rod signal is additive and depends on the luminance of the signal. As a result, photoreceptor responses can be further transformed to LMS domain by matrix $M_c(Y)$. We can here assume that two colors at luminance Y and $\tilde{Y}$ appear the same if there scaled cone response values are equal. Then the retargeted color values can be computed by:

$$\begin{bmatrix}\tilde{R}\\ \tilde{G}\\ \tilde{B}\end{bmatrix} = \frac{\tilde{Y}}{Y}\left(M_c(\tilde{Y})M_E\right)^{-1} M_c(Y) M_E \begin{bmatrix}R\\ G\\ B\end{bmatrix}$$

After the hue shift is corrected, the loss of color saturation is adjusted by:

$$\hat{R} = \left(\frac{\tilde{R}}{\tilde{Y}}\right)^{\frac{s(Y)}{s(\tilde{Y})}} \tilde{Y}. \text{ where } s(Y) = \frac{Y}{Y+0.108}$$

to account for decreasing sensitivity of the cones as well as changes introduced by the tone-curve.

An example of original image and its enhanced images after each step are shown in figure 4.

3.2. Perceptual anchoring

This algorithm combines four main parts: device characteristic modeling, forward CIECAM02 model, backward CIECAM02 model and post gamut mapping. Figure 5 illustrates the block diagram of this algorithm.

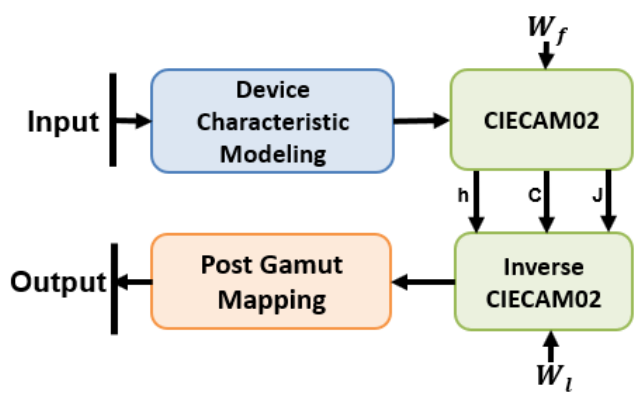

Figure 5: block diagram of perceptual anchoring method.

3.2.1 Device characteristic modeling

Device characteristic modeling involves converting pixels from srgb to lrgb domain. Since this algorithm targets faithful color reproduction, different gamma values are used for different channels for precise accuracy. Next, the input pixel value is transferred from the device-dependent RGB space to the device-independent XYZ space using full-backlight display property.

For one display, there are twelve parameters that needs to be determined in this part: three gammas for different channels and nine for RGB to XYZ transformation matrix. These nine parameters are estimated by convex optimization. The parameter estimation only needs to be conducted once until the display is recalibrated.

3.2.2 Forward CIECAM02

CIECAM02 takes not only stimulus itself, but also the white point in the display, adapting luminance, relative luminance of the background, surrounding condition into account to the model. There are five steps in the model to transfer the input to the output perceptual attributes. The flowchart of the model is shown in figure 6.

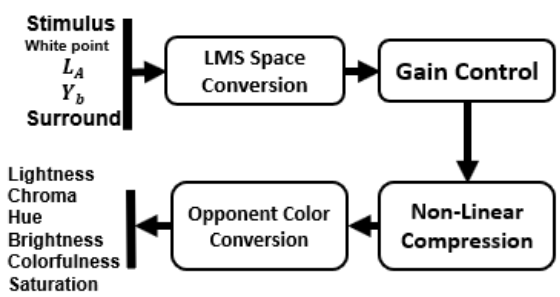

Figure 6: flow chart of CIECAM02 model.

The first step is basically to convert XYZ space to LMS using Mcat02 matrix. LMS values then are passed to the second step where they would be adjusted based on how well different types of cones can adapt to the adopted white point, and the degree of adaption is a function of the surround and the adapting luminance. The post-adaption non-linear response compression is then applied to the result of gain control. Similar to gain control, the non-linear compression is carried out in different channels for better accuracy. These first three steps together are called chromatic transform.

Opponent color conversion transfers pixels from LMS cone-response space to YCbCr domain. Then six perceptual attributes can be computed following the formula CIECAM02 model proposed. In this study, only hue, lightness, and chroma are used.

3.2.3 Backward CIECAM02

Since the goal of this study is to preserve the perceptual attributes in low-backlight display, now hue, lightness, chroma given by forward CIECAM02, and the white point in low-backlight display serve as input, the corresponding enhanced RGB values can be computed following the steps of CIECAM02, but in a reverse order.

3.2.4 Post gamut mapping

Some enhanced pixels are not displayable by the low-backlight display. Instead of hard clipping, the gamut mapped pixels are computed by blending the clipped pixel value with the original pixel value:

$$\begin{bmatrix}\tilde{R}\\ \tilde{G}\\ \tilde{B}\end{bmatrix} = (1-JC)\begin{bmatrix}R_c\\ G_c\\ B_c\end{bmatrix} + JC\begin{bmatrix}R_i\\ G_i\\ B_i\end{bmatrix},$$

where J and C represents lightness and chroma of that pixel respectively.

The same image is processed by this algorithm, and compared in figure 7.

Figure 7: result comparison of using perceptual anchoring method. The top image is the original image and the bottom image is enhance image using perceptual anchoring.

4. Results and discussion

Three images in srgb color are processed by both algorithms and the results are compared in this section. Firstly, a qualitative comparison is made and then a quantitative metric is introduced to further compare the two algorithms. In order to simulate the effect of low luminance but still put the original image on the same screen, the processed images are artificially dimmed to simulate the effect of being

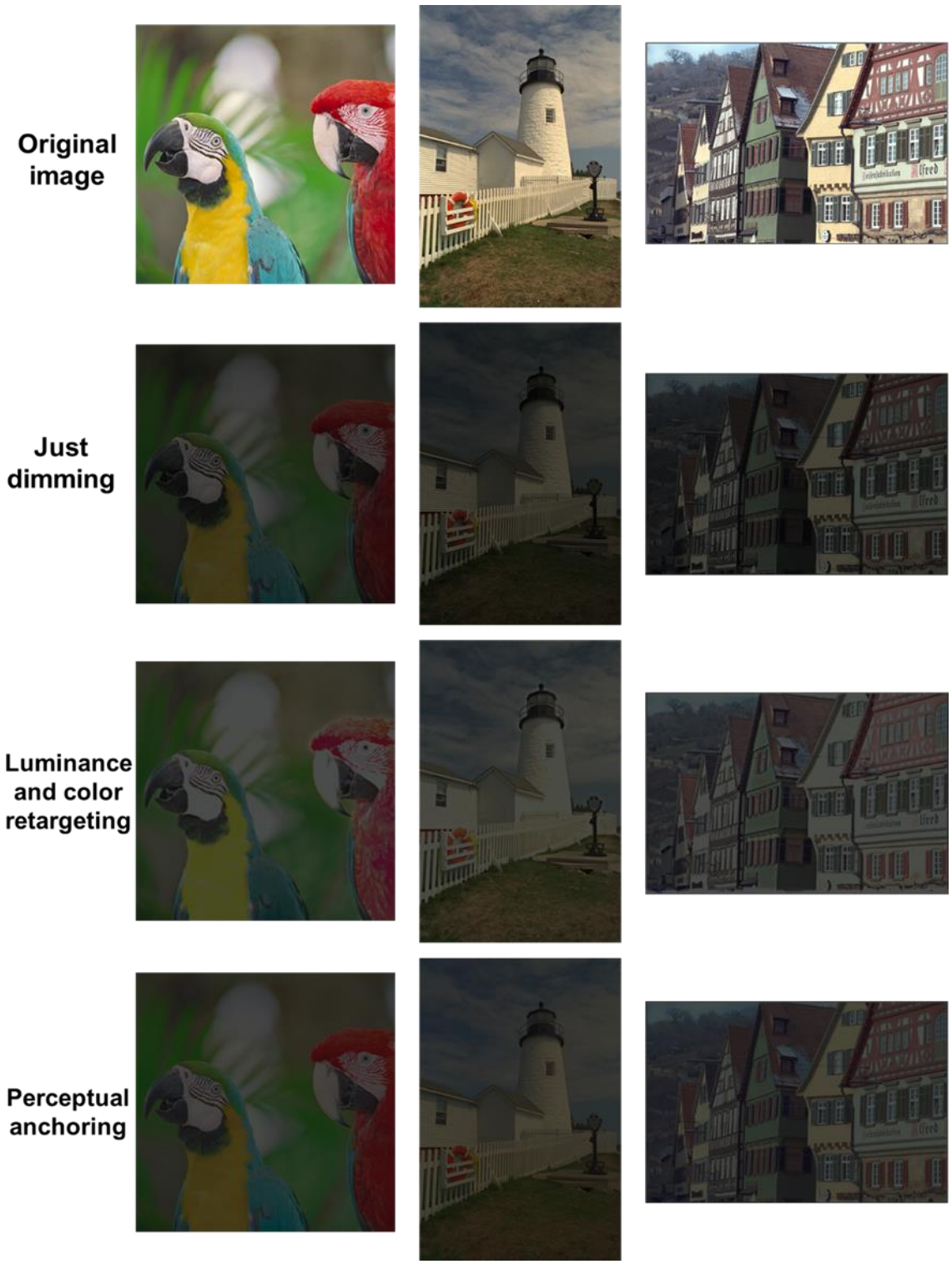

Figure 8: results comparison. The first row contains the original images, second row shows the original image with low luminance, third row shows images first enhanced using luminance and color retargeting and then dimmed, the last row shows the images enhance using perceptual anchoring method and then dimmed.

displayed on a low light display. The original images are also dimmed without modification and shown together to serve as a control and reference. The original and processed images are shown in figure 8.

It is noticeable that the algorithms do make a perceptual difference in terms of viewing experience. For example, it is easier to resolve the grass texture in the lighthouse image when processed by contrast and color retargeting. However, in the case of the bird image, the texture of the red feathers is over enhanced. As a result, it cannot be certain that one method is superior than the other, the situation is image specific. According to our implementation of the two algorithms, there is a significant difference in their computational complexity. This is reflected by their process time. Both algorithms are executed on a laptop with Intel i5-3230M processor with 2.60 GHz. To process the buildings image, the perceptual anchoring method took 0.97s to finish, while the contrast and color retargeting method took 16.2 minutes, although the original paper reported 5 seconds average time to process a single HDR image on an Intel i7 CPU [3]. After examining the timing budget, the most time-consuming task is computing the CSF, which took 961 seconds. If these functions can be precomputed and use a look up function when needed, a lot of time can be saved.

In order to quantitatively compare two images in their perceptive differences, a metric is needed. It needs to accommodate for the different contrast and color response functions of the eye under different luminance. In this sense, peak signal-to-noise ratio (PSNR) is not useful, even it is commonly used to compare image differences. In this work, a visual metric that takes luminance and contrast sensitivity functions into consideration [10] is used to compare the results obtained by the two methods. The visual difference predictor for HDR images (HDR-VDP-2) is a calibrated visual metric that compares a pair of images and predicts their visual differences and quality degradation. To make such a comparison that emphasizes on the luminance of the images. The images are first converted into YCbCr space and only Y channel is extracted. This is because HDR-VDP-2 only accepts images with only one color channel that contains absolute luminance values provided in cd/m^2 [10]. The Y channel of the reference image is scaled to a maximum of 100 to represent the condition of 100 cd/m^2 viewing condition. The processed images are scaled to 1 to represent the condition of being dimmed down to 1 cd/m^2. Each of the processed image is then paired with the reference and compared by the hdr-vdp-2 algorithm. Three images are processed and compared. The results are shown in figure 9. Hdr-vdp-2 gives the difference map as well as a quality factor Q. Although the absolute luminance is considered, the differences for both methods are very high. The values in the difference map represent the probability that the difference between the images are visible for an

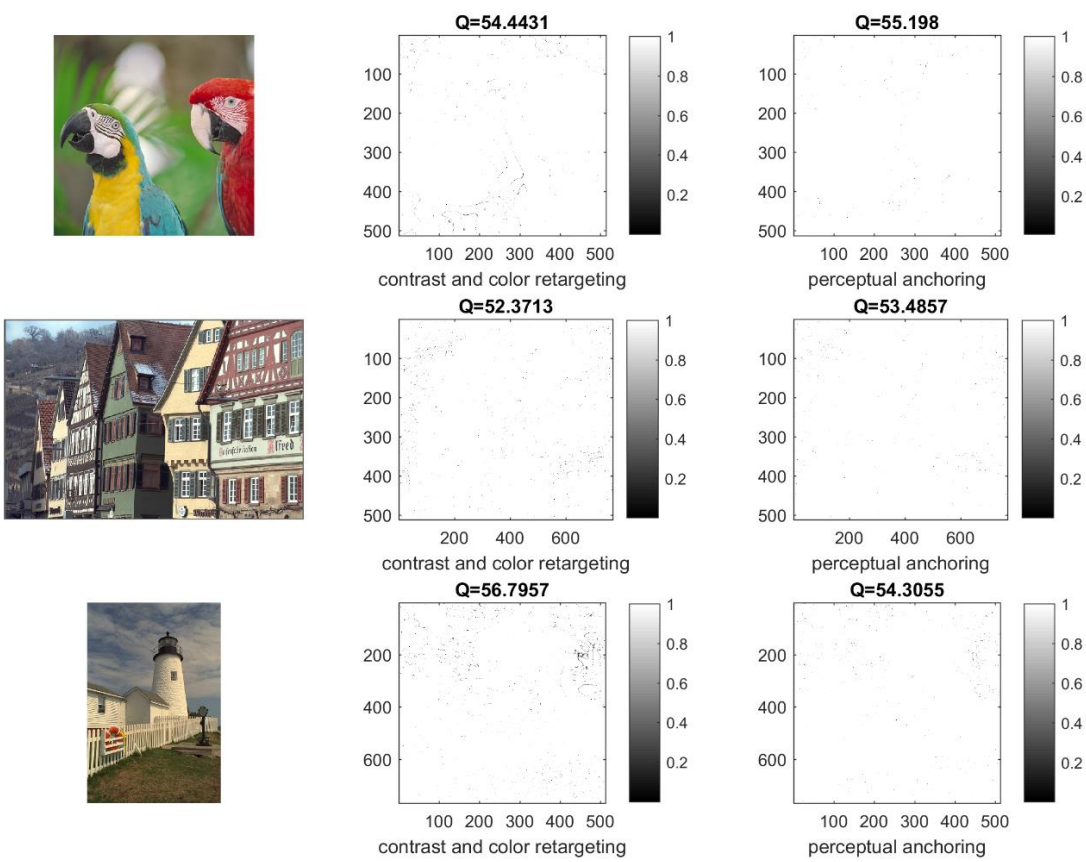

Figure 9: comparing results of the two enhancing algorithms using HDR-VDP-2.

average observer, with 0.5 as the threshold. It can be observed that either way produces a very high probability of distinguishing. This is not too surprising because even though the luminance is taken into consideration in shaping the contrast sensitivity function, difference luminance values still make a difference to the viewer. A small probability of distinguishing in this situation is not realistic. However, it is still useful to look at the quality factor Q. This measures the quality degradation with respect to the reference image, with 100 being the best quality and gets lower for lower quality. It is noted that not all images have a better quality when processed and dimmed using the contrast and color retargeting method. For example, the bird image and the lighthouse image have opposite quality comparisons.

5. conclusion

When the brightness of the display is reduced due to the low ambient light, human eyes have a different set of response functions. Therefore, the same content would appear different. In order to create a consistent image perception, the image needs to be adjusted. In this work, we have investigated two different algorithms, namely, contrast and color retargeting and perceptual anchoring. The two methods each creates an image that is more pleasing to the eye. The visual difference is mainly reflected by the color saturation and high spatial frequency components. The contrast and color retargeting method is much more computationally complicated than the perceptual anchoring method, since it treats different frequency bands separately. It can be concluded that certain images work better with the contrast and color retargeting method and others work better with the perceptual anchoring method. Since the contrast and color retargeting method emphasizes heavily on the contrast sensitivity functions, it is predicted that images with a higher range of spatial frequency will work better with this method. To verify this prediction, more images need to be processed and compared. In addition to this, a different comparison metric could be investigated and developed to better capture the visual difference with difference luminance levels.

References

[1] Moroney, Nathan et al. 2002. The CIECAM02 color appearance model. IS&T/SID Tenth Color Imaging Conference.

[2] Shih et al. 2016. Exploiting perceptual anchoring for color image enhancement. IEEE Transactions on Multimedia.

[3] Wanat and Mantiuk. 2014. Simulating and compensating changes in appearance between day and night vision, Proceedings of ACM SIGGRAPH 2014. ACM Trans. Graph. 33, 4.

[4] BARTEN, P. G. J. 1999. Contrast sensitivity of the human eye and its effects on image quality. SPIE Press.

[5] Mantiuk et al. 2013. Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance. IS&T/SPIE Electronic Imaging.

[6] Sullivan et al. 1975. Contrast constancy: deblurring in human vision by spatial frequency channels. The Journal of Physiology 252, 627–656.

[7] BARTEN, P. G. J. 1999. Contrast sensitivity of the human eye and its effects on image quality. SPIE Press.

[8] Brady et al. 1995. What's constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns. Vision Research 35, 6, 739–756.

[9] Cao et al. 2008. Rod contributions to color perception: linear with rod contrast. Vision Research 48, 26, 2586-92.

[10] Mantiuk 2011. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. ACM Trans Graph 30, 4, 40:1–40:14.