

Image Dependent Gamma Selection Based on Color Palette Equalization and a Simple Lightness Model

Juha Katajamäki and Pekka Laihanen
Laboratory of Media Technology, Helsinki University of Technology
Helsinki, Finland

Abstract

A method for estimating the best gamma value for an RGB image is described. The algorithm is intended to be part of a larger image enhancement system. It first quantizes the image colors into a small palette and then finds the gamma that distributes the palette colors most evenly in a modeled perceptual lightness scale, which adapts to the overall image luminance. The algorithm was tested both against gammas assumed to be correct and subjectively optimal gammas. The estimated gamma offers no statistical advantage when a reliable correct gamma is known, but when one is not available, the algorithm can be used to select gammas that correlate with the subjective optimum to a degree comparable to differences between individual subjects.

Introduction

A fundamental part of most applications where RGB images are used is the definition of the gamma value. In this paper, gamma means the exponent of the power function describing the transformation from RGB signal to intensity, i.e. larger gamma implies darker appearance. Wrong gamma may result in a poor contrast and color balance as well as a wrong overall lightness level. Moreover, it is very difficult to correct these deficiencies with other color adjustments. Thus it is very important to initially select the right gamma for all images. If this value is not known, it cannot be presumed that any logically chosen default value (such as 2.2) would give good results regardless of image type and origin.

Gamma selection has been an important step in the automatic color image enhancement algorithms developed at the Helsinki University of Technology (HUT) in the recent past.[1] This paper discusses the latest gamma selection algorithm developed at HUT. The primary goal of this algorithm is to find appropriate gamma for images coming from different sources (e.g. 2.2 for “PC-images”, 1.8 for “Mac-images”, 1.0 for certain scanners used in graphic arts applications etc.). In addition, gamma (together with other adjustments) is used as a general adjustment tool, which may be needed even when a predefined gamma value is known.

In this paper all the tests were made using CRT monitor images. The gamma algorithm can, however, be used for other types of images too. It is planned to be the first step in automatic color enhancement procedures which may or may not include output device dependent color adjustments.

Selection of Gamma

The basic idea of the gamma algorithm is the same as in histogram equalization: to distribute lightness levels evenly between black and white. Instead of accepting general transformations in color space, the transformation is constrained to gamma correction, and the gamma value is sought which gives the most even distribution. Besides this constraint, there are two essential differences compared to basic histogram equalization. First, instead of using the histogram, the RGB image is converted to an indexed image with few colors and those colors are distributed evenly. Second, the lightness is expressed in a perceptually uniform scale, which adapts to the overall brightness of the image. Hence the selected gamma may be different for images with the same color palette but with different frequencies for the colors. Similar adaptation schemes have been used in the context of grayscale histogram modification.[2]

Palette Generation

A well-known drawback of simple histogram equalization is over-enhancement of contrast in spatially smooth regions. As an example, for images with a distinct object against a smooth background, the equalization is affected very much by the size of the background. Using a color palette, we assume we can give the image colors visually more relevant weights than by using a histogram. The method used for the indexed image creation is a clustering algorithm that first generates quantization centers in the color space, then classifies each pixel according to the nearest-neighbor rule. The partition of the color space thus defined is not necessarily equiprobable. This is the essential advantage over histogram equalization; it means that even large smooth regions are usually represented by just a few colors, whose weight in gamma optimization is the same as the weight of other colors.

Modeling of Perceived Lightness

Although the color palette is a much better basis for gamma selection than a histogram, it is not sufficiently adaptive to the image content. To improve the method, some kind of adaption scheme is needed. Unfortunately, the restrictions of current color science and the need for a relatively simple, computationally efficient solution, mean that image dependent perceived lightness can only be estimated with very limited exactness. This does not, however, mean that image content should be ignored completely.

Tests with different photographic color images showed that the average lightness level of the darkest images was about 10 CIE L* units while the level of the lightest images was 80. When small patches of different shades of gray were placed on a dark and a light image the lightness difference ΔL* between patches producing the same visual perception could easily exceed 40. This result encouraged efforts to find a simple method that would at least partially take into account the image content when measures of lightness are calculated. The approach chosen was based on the empirical fact that visual effects of simple uniform backgrounds can be predicted relatively accurately. An assumption was made that a color image affects the perception of its colors in roughly the same way as would a uniform background with the average color of the image. Furthermore, it was assumed that if this is the case, using the average color of an image as the adaptive reference would be beneficial when estimating perceived lightness.

In practice, the key question is, how much the accuracy of this assumption is affected by the image content of normal photographic images (how badly it fails?). This was tested visually by comparing small gray objects (less than two degrees; the same shape and size but varied tone) placed on corresponding locations of a selected color image and a uniform background having the average color of the image.

Visual comparison showed that the lightness difference ΔL* between patches with the same visual appearance was usually relatively small. This difference was naturally dependent on the image and the image location. In exceptional cases, at some lightness levels and image locations, ΔL* exceeded 20 but for most images and locations the differences were much smaller. Regardless of the lightness level tested ΔL* of 5 was rarely surpassed and typical ΔL* values were at an acceptable level of two or even lower.

These results suggested that using image dependent average color as adaptive reference when estimating perceived lightness is not an accurate but nevertheless a useful approach in practical applications where simplicity and computational efficiency are important.

This adaptation scheme proved to be particularly beneficial in cases of inherently dark images whose gamma values tended to be too small otherwise.

Given a gamma value γ, the lightness is defined as follows. First, the Y tristimulus value of an RGB color is calculated, assuming sRGB primaries:

$$Y = 0.2126\, R^{\gamma} + 0.7152\, G^{\gamma} + 0.0722\, B^{\gamma} \qquad (1)$$

Then, a nonlinearity similar to the CIELAB or CIELUV color spaces is applied:

$$L = f(Y/Y_a)/f(Y_{ref}/Y_a), \quad \text{where } Y_{ref} = 1 \text{ and}$$

$$f(t) = \begin{cases} (100 + (p-1)\,f_b)\,t^{1/p} - (p-1)\,f_b, & t > t_b \\ (f_b/t_b)\,t, & t \le t_b \end{cases} \qquad (2)$$

$$\text{with } t_b = (p f_b\,/\,(100 + (p-1)\,f_b))^{p}$$

The parameters p and f_b are fixed and determine the general form of the function: p is the inverse exponent and f_b is the endpoint value of the linear portion near black (for CIE L*, $p = 3$ and $f_b = 8$). The Y_a adaptation parameter is related to the average Y value of the image. We have used the simple relation

$$Y_a = Y_{ave}/b \qquad (3)$$

where b is a constant smaller than unity. When Y_{ave} increases, function f becomes less nonlinear. This has the effect of shifting the emphasis in lightness differences from the dark end towards the bright end, which forces the algorithm to select a smaller gamma. The division by $f(Y_{ref}/Y_a)$ normalizes the lightness into [0, 1]. It is the short linear portion that makes the normalized L vs. Y curves dependent on the adaptation. If f were just a power function, the normalization would cancel the adaptation effect completely.

This method in its present form is mainly a mathematical trick having an effect in the right direction. Anyway, its similarity to visual phenomena, as modeled e.g. by CIECAM, gives it some additional justification. We have also tested a CIECAM97s based version and obtained comparable results.

Dynamic Range Normalization

If the RGB range of the image does not extend to the maximum available signal value due to underexposure or similar reason, equalizing the original colors would often result in a small gamma. Therefore, the algorithm does not use original RGB values in (1) but scales them by a dynamic range estimate RGB_{max}:

$$(R, G, B) = (R_{orig}, G_{orig}, B_{orig})\,/\,RGB_{max} \qquad (4)$$

The Algorithm

The entire algorithm consists of the following stages:

1. As preprocessing steps, downsample the RGB image if it is very large, then process it with a nonlinear filter to remove noise and small details.
2. Generate an RGB color palette with a clustering algorithm and compute the pixel counts for the palette colors. Let the palette size be M, and let the color vectors be denoted by $\mathbf{c}_i$ and the counts by N_i, $i = 1, ..., M$.

3. Find a dynamic range estimate RGB_{max} from the image histograms. Scale the palette colors using (4).
4. Divide the normalized lightness interval [0, 1] into M parts of equal length and take their centers as target lightnesses L^t_1, ..., L^t_M, where L^t_1 is the smallest lightness etc.
5. Repeat for $\gamma = \gamma_{min}$ to γ_{max}:

5a. Get the Y components of the $\mathbf{c}_i$ using (1). Denote these by Y_i.

5b. Compute Y_{ave} as the weighted mean of Y_i, using weights N_i. This is approximately the same as the average Y over all pixels. Define the adaptation level according to (3).

5c. Get the visual lightness values L_i using (2).

5d. Sort the L_i. For each k = 1, ..., M, let $i(k)$ be the index of the palette color whose lightness is the kth smallest. Calculate the sum of squared errors between the actual and the target lightnesses:

$$e(\gamma) = \sum_{k=1}^{M} \left(L^t{}_k - L_{i(k)} \right)^2 \qquad (5)$$

6. Select the gamma for which the error is smallest:

$$\gamma_{opt} = \arg\min_{\gamma} e(\gamma) \qquad (6)$$

Invariance Properties

A desirable property of a gamma selection algorithm is invariance to changes in target gamma. In other words, the automatically selected gamma should vary with the image in a way that cancels gamma corrections. If an image I_1 is transformed into I_2 by raising its pixel values to the power of γ, and if γ_1 and γ_2 are the optimal gammas selected by the algorithm for I_1 and I_2, respectively, then it should hold: $\gamma_2 = \gamma_1/\gamma$.

The invariance holds exactly if the indexed image does not depend on gamma, and if the palette colors follow the same gamma transform as the image itself. This is true for ideal images consisting of only a small number of objects with constant and sufficiently distinct colors, as long as the gamma stays within certain limits. Therefore, the invariance is expected to hold well for real images whose color distribution is very strongly concentrated into a small number of clusters. To be gamma invariant in general, the algorithm should optimize some criterion that is a function of the whole image, not just of the color palette. This would imply higher computational complexity, so the palette approach was adopted as a compromise.

Under the above ideal assumptions the algorithm is, due to the RGB normalization (4), also invariant to linear scaling of RGB values. That is, the same gamma is selected for I_1 and I_2 if $I_2 = a\, I_1$ with a constant. For real images, this invariance is only approximate because the palette generation uses the original unnormalized image.

Results

To test the algorithm, we selected 300 pictures portraying various subjects, including people, landscapes, buildings, interiors, nature, close-ups of objects, etc. Different times of year and day were represented. The pictures were on Kodak Photo CDs, from which they were read into Adobe Photoshop in CIELAB format. The pictures were converted to RGB using Photoshop settings that corresponded to the sRGB space except in the gamma, which was varied so that several gamma versions could be generated from the same picture. The gamma values used were 1.0, 1.4, 1.8, and 2.2.

The picture set was divided into 100 training pictures and 200 test pictures. The test set had not been used previously during development of the algorithm.

Estimated Gamma Versus Assumed Gamma

The algorithm was first tested against assumed target gammas. To optimize the parameters, all four gamma versions of each training picture were generated. Preliminary tests with some parameter combinations had showed that several combinations led to statistically similar results. Therefore, in the tests described here, parameters p and b in Eqs. (2) and (3) were fixed at the values 3 and 0.3, respectively, and only f_b was varied. The root-mean-square error for all 400 training images between the logarithms of the estimated and the assumed gammas was computed with different values of f_b. The optimal value of f_b was found to be 12. In all subsequent comparisons, as well as above, the logarithm of the gamma rather than the gamma itself is used because the log scale is visually more uniform and because, due to the invariance property discussed above, variation in the estimated gamma is expected to be proportional to the average gamma.

The linear portion of the resulting lightness function is somewhat longer than that of the CIELAB curve, causing the curves to depend more on the adaptation parameter Y_a. Figure 1 shows examples of the resulting curves with different adaptation parameters.

Four gamma versions of the test images were also generated as above and the algorithm was run for this set, whose total size is 800. Table 1 contains some statistics on the results both for the training and the test images and for each fixed reference gamma. The overall RMS error in the natural logarithm domain is also shown in the table for the training and test sets. For example, the log-domain error value 0.27 corresponds to a multiplicative error of exp(0.27) = 1.31 in the gamma domain.

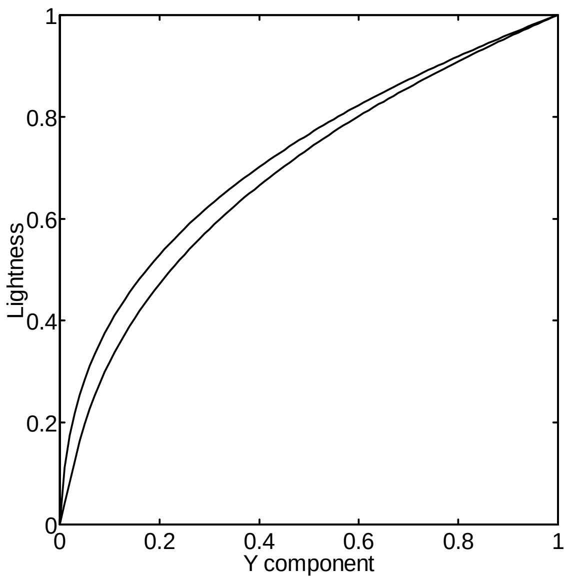


Figure 1. Model of Eqs. (2) and (3) for perceptual lightness vs. Y component at different adaptation levels. Lower curve: Y_{ave} = 0.4. Upper curve: Y_{ave} = 0.05.

Table 1. Comparison of algorithm results with reference gammas.

Reference gamma	1.0	1.4	1.8	2.2	All
Mean (training set)	1.23	1.49	1.78	2.01	
Coef. of variation (training set)	0.25	0.28	0.29	0.29	
RMS error in log domain (training set)	0.29	0.24	0.24	0.28	0.26
Mean (test set)	1.20	1.47	1.76	2.01	
Coef. of variation (test set)	0.26	0.28	0.29	0.29	
RMS error in log domain (test set)	0.28	0.25	0.26	0.29	0.27

Table 2. Behavior of algorithm results as the gamma is varied. Note that the geometric mean of the ratio is equivalent to the mean difference in log domain.

Reference gamma 1	1.0	1.4	1.8	1.0	1.4	1.0
Reference gamma 2	1.4	1.8	2.2	1.8	2.2	2.2
Expected γ_2 / γ_1	1.40	1.29	1.22	1.80	1.57	2.20
Geom. mean of γ_2 / γ_1 (training set)	1.21	1.19	1.13	1.44	1.35	1.62
Correlation between γ_2 and γ_1 in log domain (training set)	0.90	0.95	0.96	0.81	0.91	0.75
Geom. mean of γ_2 / γ_1 (test set)	1.22	1.20	1.14	1.46	1.37	1.67
Correlation between γ_2 and γ_1 in log domain (test set)	0.94	0.97	0.96	0.91	0.93	0.85

Table 2 describes the behavior of the algorithm for different versions of the same image. For example, when the reference gamma changes from 1.4 to 2.2, the result should change by a constant ratio of 2.2 / 1.4 = 1.57 in the ideal case. In reality, the average ratio is somewhat smaller (1.35). The correlation coefficient between the two estimated gammas is rather high (0.91 to 0.93). In other words, despite the fairly large variance in the estimated gammas, the estimation result mostly changes consistently with the reference gamma when versions of the same picture are compared, i.e. the invariance holds with reasonable accuracy.

Estimated Gamma Versus Visual Optimum

The ultimate performance criterion of the gamma algorithm is how it contributes to improving the visual quality of pictures. This was studied with subjective testing. To make the tests practical, they were conducted in three sessions using 100 pictures in each: the training set comprised one part and the test set was divided into two equally large subsets. Only one gamma version of each picture was chosen for the tests randomly, but in such a way that there was an equal number of samples (25) from each reference gamma within each of the three subsets.

During a test session, the 100 images were displayed on a monitor in random order. The subject could adjust the image appearance with a slider that controlled the gamma, and he or she was asked to set the slider at the optimum position. The picture was displayed in grayscale instead of color. This was considered the most appropriate method because the algorithm tries to optimize lightness only, without regard to saturation or color balance. Prior tests had suggested that since visual colorfulness usually increases as gamma increases, the optimal gamma tends to be higher in the presence of color information, leading to too dark images compared with the case where there is no color information. Testing of the algorithm without color was justifiable because the algorithm was not intended to be executed in isolation, but rather was designed to be accompanied by a saturation adjustment, which can compensate for the lack of saturation resulting from smaller gammas. By omitting color, we also eliminated some ‘feel’ aspects of the image which can affect the optimal gamma. Since we cannot expect an algorithm as simple as this to recognize such effects, eliminating them can even be considered advantageous for testing purposes.

For the visual optima to be comparable to the estimated gammas, the dynamic range normalization (4) was also applied to the displayed images. Given a gamma value determined by the position of the slider, the test program calculated the Y tristimulus component of the color image using this gamma and converted it into a monitor-corrected grayscale image using a measured tone reproduction curve. The gamma scale represented by the slider consisted of discrete steps that were equally spaced in the log domain. The initial slider position was chosen at random for each image.

Each of the three tests was executed twice by three subjects. One subject was one of the authors, but the other two had no background in color image processing.

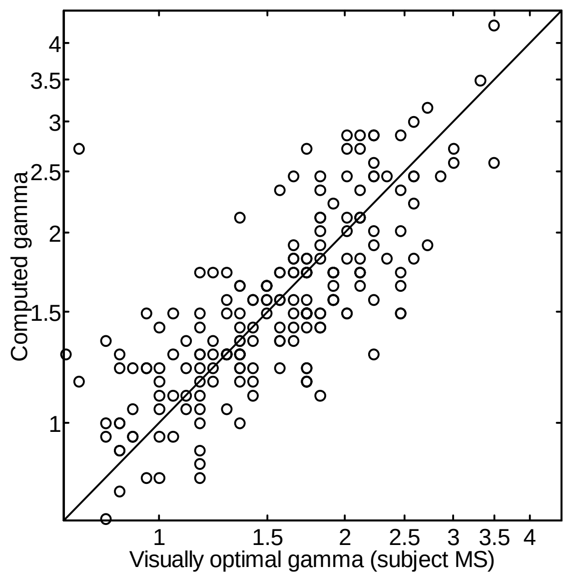
When examining the results, we must consider the possibility of a systematic difference occurring between the subjectively optimal gamma and the assumed gamma, due to viewing conditions or monitor calibration error. This difference was checked for the training images and was found to correspond to a relative gamma of about 1.03. Since this value is close to unity it was ignored, and the subjective gammas were used in the comparisons directly without any correction.

Table 3 shows the results of comparisons between the assumed, estimated, and subjective gammas for the test set. We employed a simple RMS criterion, which does not take into account the fact that the limits for visually acceptable gamma can be looser for some images than for others. Although the small number of subjects and the restricted set of images do not allow us to draw statistically reliable conclusions, the following observations can be made from the table data. First, the difference between the assumed and subjective gammas is rather large, even for the average subject. This suggests that the assumed gamma is not a very good basis for studying algorithm performance. Second, the estimated gamma is closer to the subjective gamma than to the assumed gamma. Thus, the fairly large variation in Table 1 can be partly explained by the fact that the assumed gamma is not visually optimal for all images. Third, there are significant differences between two subjects which can be of the same order as the error between the estimated and subjective gammas.

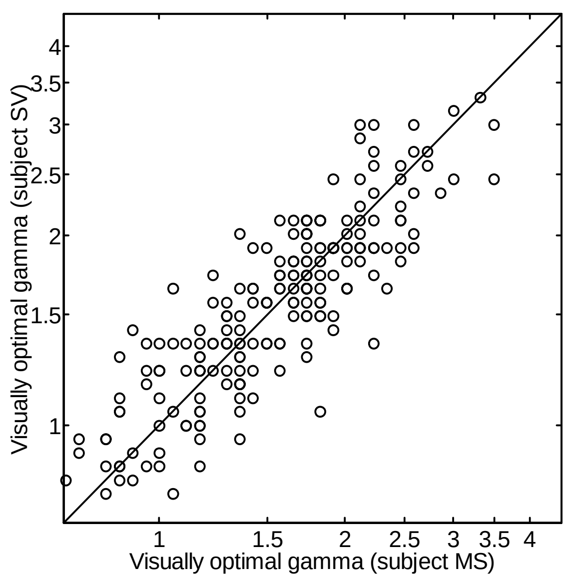
In Figure 2, examples of the correlation between the computed and subjective gammas as well as between two subjects are displayed graphically.

Table 3. Comparison between assumed, estimated, and subjective gammas for the test set (200 images). The values in the last column are exponent functions of those in the middle column. The results for individual subjects are averages of the two test sessions.

	RMS diff. in log domain	Difference in multiplicative form
Estimated / Assumed	0.28	1.33
Estimated / Average subject	0.20	1.22
Estimated / Subject JK	0.23	1.26
Estimated / Subject MS	0.23	1.26
Estimated / Subject SV	0.22	1.25
Assumed / Average subject	0.19	1.21
Assumed / Subject JK	0.22	1.24
Assumed / Subject MS	0.23	1.26
Assumed / Subject SV	0.21	1.23
Subject JK / Subject MS	0.21	1.24
Subject MS / Subject SV	0.18	1.19
Subject SV / Subject JK	0.16	1.17



(a)



(b)

Figure 2. Log-scale comparison of estimated and subjectively optimal gammas for the test set. Note that some circles may represent several overlapping points.

Further Development of Gamma Algorithms

Although the performance of the algorithm in conjunction with other corrections is still an open question, the present version is evidently too simple for certain cases. For example, optimization of just the lightness component causes over-enhancement in regions having roughly constant lightness but varying color, since such regions are represented by several palette colors. The outlier point in the upper left region of Fig. 2 (a) is such a case. Also, the adaptation to overall brightness is not quite adequate.

Usually, there exists *a priori* knowledge of the gamma lying between, say, 1.0 and 2.5. If additional gamma selection criteria, such as saturation, are used, they should be designed so as to counterbalance the lightness criterion in a manner that forces the optimum to be within the correct range.

The algorithm described above does not use spatial information, except for the small effect of the pre-filtering. Actually, the method was designed to be a preprocessing stage of a larger automatic image processing system whose aim is to utilize spatial information as well. The result of spatial segmentation corresponds to actual objects and backgrounds even better than palette colors do, so the gamma algorithm is expected to be improved if it is modified to exploit segment information, possibly with a more accurate model of perceived lightness.

The gamma, although very useful, is not the only parameter for color image enhancement. To find the best color reproduction, several parameters should be optimized at the same time. Of course, when the number of variables grows, it becomes more difficult to define an objective function whose optimum at least roughly corresponds to visual optimum. Nevertheless, the color palette approach could lend itself to somewhat more complex color adjustments as well. If the optimization involves iterative evaluation of global color properties, these can be approximated using the palette colors and their frequencies without the need to scan the image pixels repeatedly.

References

1. P. Laihanen, Printer-Specific Automatic Enhancement of Photographic Color Images. *IS&T's 1998 PICS Conf.*, pp. 405–408.
2. D. T. Cobra, Image Histogram Modification Based on a New Model of the Visual System Nonlinearity, *J. Electronic Imaging,* **7,** 807 (1998).