Image Color Transfer with Naturalness Constraints

Xuemei Zhang, Apple Computer, Cupertino, CA; Hui Chao, Qualcomm Research, Cupertino, CA; Daniel Tretter, Hewlett Packard Laboratories, Palo Alto, CA

Abstract

In consumer imaging applications involving photo collages or composition of user photos with professional artwork, inconsistent color appearance of photos and artwork from different sources can result in compositions that do no look aesthetically pleasing. Users often express a desire to modify individual images to achieve a more consistent color appearance. Prior work in color transfer that extract the color properties of one image and apply it to another have shown very interesting results [1,2]. These works focused on achieving an artistic effect, usually without the constraint of conserving object color. In consumer imaging, we have to be more conscious about conserving general object color and especially skin tones, which are not amenable to aggressive color change. In this paper we describe an algorithm to estimate the color and tone properties of an image and transfer these properties to another image under a strong naturalness constraint. In our method, color changes are constrained to correspond to incomplete adaptation under natural illuminants. We use a simple Bayesian method to characterize scene color properties, expressed as scene color temperature and illumination levels. An existing color adaptation model RLAB [3] is used to apply color changes by simulating incomplete adaptation to a colored target illuminant. We emphasize that this is not a method of white point estimation nor a white balance procedure Rather, we use color adaptation models as a means to ensure color adjustments to be "plausible", and therefore maintain a natural appearance to the images even after significant color adjustments.

Introduction

As personal publishing starts to replace simple photo printing as the dominant consumer photo activity, there is an increasing demand for automatic composition of multiple photos into a themed page with a professionally designed appearance. Photos from user collections are frequently from several digital cameras of varying qualities, taken at different time, locations, and events, and occasionally from scanned materials or images found on the internet. Users are rarely willing to separate related content or group unrelated content simply based on color consistency. Therefore, it is not uncommon to see images of different color characteristics and tone properties being put into the same composition due to their content relevancy (Figure 1). In addition, most photo composition softwares provide professional template artworks for a page, which tend to have very different style and color characteristics from typical consumer photos. To create an attractive page composition in these circumstances, it is important to provide the user with tools that automatically adjust image color properties to be consistent with other images and template artwork in the composition.

***Figure 1**. Images with related content but different color properties that might go on the same page*

The most straightforward way to achieve color consistency in a composition is to make all the photos share similar color properties as the focal photo on the page, or as the theme artwork. In both cases this can be done as a color transfer: extracting the color properties of the focal photo or the theme artwork images, and transferring them to all photos on the page. The problem of color transfer has been studied with good results in prior work. In Reinhart et al's 2001 study [1], a target image's color is changed by adjusting the mean and standard deviations of its pixel distribution in 3 color channels of an opponent color space lαβ to be similar to the distribution of the source image. They demonstrated good color similarity and natural appearance when the source and target images share similar contents, or when both images have fairly homogenous colors. Xiao & Ma [2] extended this work to allow user selection of a swatch in source and target images, to enable more controlled transfer directions. Hou & Zhang [4] proposed a color manipulation method that can change an image's colors according to a "color concept" that is based on hue clustering on an library of natural color images. Color changes using this method can be drastic, and the target is artistic effect instead of preserving natural appearance of images. There is also a large body of work on image relighting, which mostly focused on shape and lighting geometry changes. These methods often need original lighting and 3D scene information, or sophisticated estimation of these information from the source and target images, which are not available in consumer imaging applications.

Our target application is photo composition in consumer publishing. Consumer photos tend to be dominated by faces and familiar objects. Therefore, aggressive color changes that might turn faces green or blue, or completely change a familiar object's color, are not acceptable to users. We aim to develop a color transfer method that satisfies a strong naturalness constraint, so that photos in the same composition can have reasonably similar appearance while individually still looking like real photographs to the user. We first characterize scene color properties as plausible scene color temperature and illumination levels. An existing color adaptation model RLAB (Fairchild & Berns 1996 [3]) is used to apply color changes by simulating incomplete adaptation to a daylight illuminant with the target color temperature and target luminance level. Without defining a specific naturalness metric, we can nevertheless preserve a level of naturalness by modeling color changes as incomplete adaptation to natural illuminants that may occur in a user's real viewing experience.

Characterizing image color properties

The first step of color transfer is characterizing source image colors with a limited number of parameters. In preparation for color adjustment based on incomplete illuminant adaptation, we seek to characterize input images with illumination-related properties such as illuminant color temperature (scene CCT) and scene luminance level. This procedure is not an illuminant estimation process, as most consumer images come from mid-range digital still cameras with decent white balance results. We only seek to describe the collective characteristics of an image's color with parameters that are amenable to transfer to another image through an illuminant adaptation model.

To go from a collection of image colors to an estimate of plausible illuminant condition, we use a simple Bayesian formulation, where the posterior probability of a particular illumination condition can be calculated from the likelihood of images under each illumination condition and the prior probabilities of these illumination conditions:

$$P(cct=c, lum=y \mid r_i,g_i,b_i) = P(r_i,g_i,b_i \mid c,) * P(c,y) * k, \qquad (1)$$

where k is a constant that does not affect our calculation in any meaningful way.

The likelihood function $P(r_i, g_i, b_i \mid c, y)$ is determined by simulating rendered colors of surfaces under different illuminant CCTs and luminance levels. The Vrhel surface reflectance collection [5] was used since it is reasonably representative of natural object surfaces that are commonly seen in consumer photos. For illuminants, we simulated daylight spectra at 9 color temperatures ranging from 3000 to 11000K, and at 9 different luminance levels from 1-10000 cd/m2, for a total of 81 illuminant conditions. CIE XYZ values were calculated for each surface and illuminant combination, and then adapted to the illuminant white point using the RLAB adaptation model [3], to obtain a set of appearance XYZ values, which are then converted linear sRGB values. Figure 2 shows an example of the rendered appearance of the same surface collection under two illumination conditions. It is clear that the same surfaces may have quite different but distinctive appearances under different illumination conditions even when color constancy is taken into account.

***Figure 2**. Example of rendered surface colors under different illuminant conditions.*

RLAB only gives us simulated appearance values for the 170 Vrhel surfaces, which are too sparse to get an empirical likelihood function, even with simulated noise added to the response values. We need to fit the a functional form so that a likelihood value can be calculated for any color that may appear in an image. A separable joint Cauchy distribution is fitted to the simulated sRGB values per illuminant condition. To ensure that fitting a separable 2-D Cauchy function is valid, we first transform the input sRGB values into a 2-dimensional space where the distribution is nearly separable. The coordinates of each color in this new space are represented by [v1, v2]:

$$[v_1, v_2]'_{(c,y)} = B_{(c,y)} * V_{rgb}, \qquad (2)$$

where V_{rgb} represents the XYZ coordinates of input color, and $B_{(c,y)}$ is a 3x2 transformation matrix specific to the illuminant condition (c, y) (sceneCCT = c, scene luminance = y). We now fit a separable 2-D Cauchy to the distribution of $[v_1, v_2]$ values:

$$P(r_i,g_i,b_i \mid c,y) = C(v_1, \square_{1(c,y)}, x_{1(c,y)})*C(v_2, \square_{2(c,y)}, x_{2(c,y)}), \qquad (3)$$

Equation (3) allows us to calculate the likelihood of any color under an illuminant condition (c, y). For an input image, the likelihood of all colors under each illuminant condition is calculated as a product of likelihood of individual pixel color, assuming (very roughly) independent likelihoods for each pixel:

$$L(R, G, B \mid c, y) = \prod_i P(r_i, g_i, b_i \mid c, y) \qquad (4)$$

Given (4), we only need to specify a prior distribution P(c, y) in equation (1) before posterior probability of illuminant conditions could be calculated for an image. This prior distribution should represent the base probability of each illuminant condition as best descriptors of typical consumer images. We do not have a large enough data set to estimate this distribution reliably, therefore we simply used an estimate based on our experience of what typical lighting conditions are in consumer photos. The prior distribution used in our implementation is a 2-D Gaussian in the CCT - luminance space that has a peak near 5500K and 1000 cd/m^2, with a fairly flat spread.

The posterior distribution of illuminant conditions can be calculated from an image's pixel RGB values according to Equation (1) using the likelihood function and prior distribution above. The scene CCT and luminance combination with the largest posterior probability is our descriptor of the image's colors, and the entropy of this distribution over different illuminant conditions can be used as a confidence measure of this estimate. Figure 3(a) shows a few image examples with color descriptors calculated this way, expressed as two numbers per image -- scene CCT and luminance. These values will be used to perform color transfer to and from other images. We emphasize again that these values are not estimates of image white point or scene illuminant, but are a set of descriptors used to summarize the colors based on their appearance.

Color transfer using a color adaptation model

Once we have estimated the plausible scene CCT and luminance for both the source image (image to be adjusted) and target image (image whose color property is the target of our adjustment), we perform the color adjustment of the source image by applying a color adaptation model RLAB, the same model we used to simulate image color under each illuminant condition in the previous section.

There are more updated color appearance models than RLAB, such as CIECAM02. The choice of RLAB as our adaptation model was based on two considerations. First, RLAB's adaptation calculation is a sequence of matrix multiplications that can be combined into a single matrix after the estimation stage is done, and therefore can enable a fast real-time implementation in a practical system. Second, the RLAB model has a strong luminance-dependent color adjustment element, which results in a large range of color saturation changes based on scene luminance level. This provides a means to adjust the color saturation of an image in a way that is consistent with variations people might see under different illuminant changes. This way, we can satisfy a naturalness constraint without defining a naturalness metric based on pixel values.

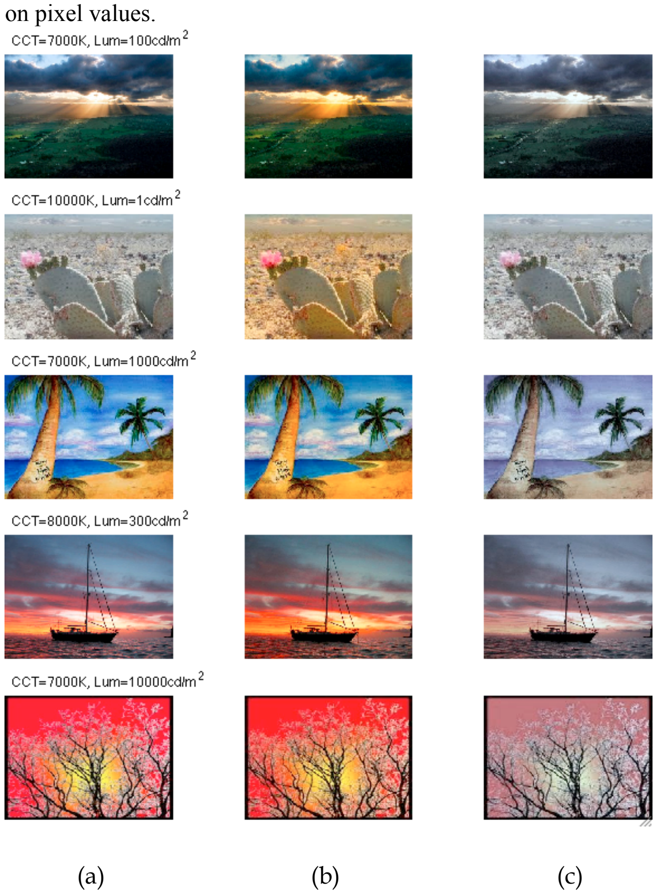


Figure 3*. Example of images with color adjustment to a specified CCT and luminance. (a) original images, (b) after color adjustment to cct=2000K and luminance=3000cd/m2, (c) after color adjustment to cct=11000K and luminance=0.1cd/m2.*

The RLAB model is described in [3] in detail. Color appearance is represented as color values in a reference condition, and appearance matching between different illuminants is done through matching their reference XYZ values. Conceptually, our color transfer is a process where we try to re-illuminate the scene with the target illuminant condition, and then rendering the scene for the same display. We need white point XYZ values for both the source and the target illuminant conditions to perform this calculation. Since illuminant conditions are described only as CCT and luminance levels, we calculate the associated white point XYZ values from simulated daylight spectra at the same CCT as the source or target scene CCT descriptor. These two white points can be represented as two sets of XYZ coordinates W1 and W2. From the source and target illuminant condition white point XYZ values, we can calculate the RLAB adaptation matrices that goes from image XYZ values to reference XYZ values. Let F(xyz, c2, y2) be the RLAB adaptation matrix of image XYZ values from source illuminant condition to the reference condition, and G(xyz, c1, y1) be the adaptation matrix from target illuminant condition to the reference condition, and H(rgb) be the transformation matrix from image linear sRGB values under display condition to XYZ values under reference condition, then our color adjustment to the target color characteristics is a sequence of matrix operations that

(1) transform image RGB values to reference XYZ values (representing their rendered appearance) using H,

(2) transform the appearance XYZ values back to scene XYZ values using F^{-1},

(3) calculate the scene XYZ values under the new illuminant condition, by independent scaling of the cone responses $M^{-1} * [(MW_2)./(MW_1)]$, where M is the conversion matrix from XYZ values to cone responses,

(4) transform the new scene XYZ values to reference XYZ values using G, and then render for a display using H^{-1}.

This whole sequence can be combined into a single transformation matrix T:

$$T = H^{-1} * M^{-1} * [(MW_2)./(MW_1)] * F^{-1} * H, \qquad (5)$$

which can be calculated once for each source and target image pair, and then applied rapidly to each pixel of an image to obtain the color-adjusted image very efficiently in a real application.

To make several images in the same composition share common color properties, we can choose a "focal" image as the target, and adjust colors of all other images toward this target. The use of scene CCT and luminance values as color descriptors also makes it easy to adjust all images to a specified CCT and luminance combination, or to an "average" target, i.e. adjust scene CCT and luminance to the average of estimated scene CCT and luminance values of these images. The examples in Figure 3(b) and 3(c) show how this transfer process can modify color properties of several images to a common target to make them more consistent in appearance.

Characterization and transfer of image tone properties

If user photos come from different sources such as different cameras or from the internet, the images can vary in both color and tone properties. Adjusting color using our method helps bring the appearance of these images closer to each other, but sometimes color adjustment alone isn't enough. To add a "tone-transfer" on top of color transfer, we adopt a similar approach and first find a small set of parameters that describe the tone characteristics of an image. We found that using two simple statistics (mean and standard deviation) on L* values of image pixels gave good results. The use of L* statistics is partially based on results from an image tone preference study [6]. For stability consideration we use only pixels around "busy" areas in an image to calculate tone statistics. Local busy-ness index is calculated as a highly blurred edge energy map. Pixels with busy-ness index above a threshold is included in the L* statistics. The busy-ness threshold is determined by scene key -- threshold is higher for very high or very low key scenes, which ensures that large bright or dark regions do not bias the L* statistics too much, and won't be adjusted to gray in the tone transfer process. The target L* mean statistic is constrained within a threshold distance from the existing L* mean, to avoid tone adjustment too far from the original, which can cause severe distortion of image appearance and loss of spatial detail after tone adjustment for some images.

Tone adjustment is done with a simple S-shaped global tone curve for most images. However, when the L* target results in smaller dynamic range for the image, we perform a local tone mapping using a simple retinex-type algorithm to compress image dynamic range before global tone adjustment. This step is necessary to preserve spatial details through the tone transfer process.

Figure 4 shows a flow chart illustrating the steps we described so far to do color and tone transfer for images to be used in composition.

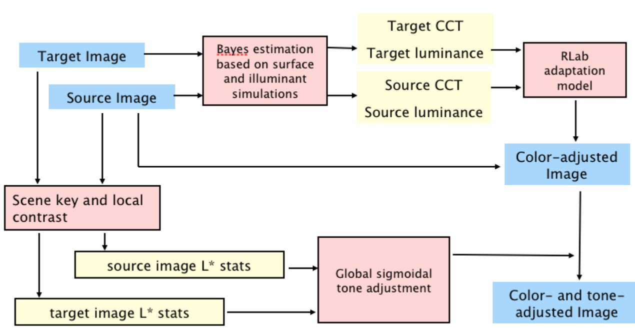


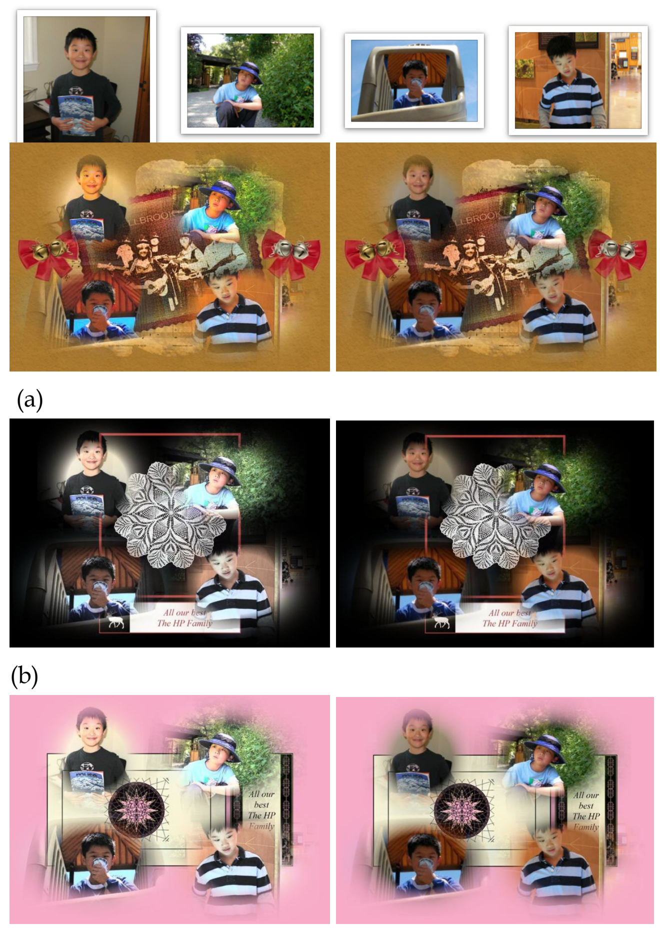
***Figure 4**. Steps in color and tone transfer with naturalness constraint.*

Results

We applied the color and tone transfer algorithms in a couple of real-world application scenarios. Adjustment of several images to a common target appearance was discussed earlier (Figure 3). Here we specifically look at applications where images are adjusted to be more harmonious with designed background and accent artwork.

In this case color and tone descriptors were calculated for both the source images and the artwork images. The artwork images are treated exactly the same way as regular photos, regardless of how they are generated. We then transfer the color and tone properties of the artwork to user images. User photos can be color- and tone-adjusted to blend well with design artwork of significant variations while still looking like "real" photos. Figure 5 shows results of composition of the same set of user photos to 3 different page background designs. The user photos (top row) are quite different from each other in both dominant color and in tone properties, and all of them different from the artwork color properties in some way. The compositions are arranged in 2 columns, the left showing compositions after our color consistency algorithm, and the right column without such adjustments.

Composition in Figure 5(a) demonstrates that our color and tone transfer process is not a simple case of forcing all colors to be similar to the artwork color -- which would result in highly unnatural colors for the sky and grass; instead, we find the scene color descriptors and then use an adaptation model to push all colors as a whole to the target descriptor. The typical result of our operation is that there will be more shared colors between target



(c)

(1) *with consistency* (2) *without consistency*

***Figure 5**. Examples of page compositions with color and tone transfer from artwork to user photos.*

image and transformed source image, but not all source colors are pushed to be similar to target colors. In figure 5(a) the sky and grass colors stayed blue and green, while skin tone and some wall colors changed to be more consistent with artwork colors.

The artwork in Figure 5(b) and 5(c) have vastly different color and tone properties, but the estimated color descriptors are fairly similar for the two of them, and thus the adjusted colors of user image are also similar between these two compositions. It is easy to see that the color and tone adjustments in both cases resulted in shared colors with the background artwork, and gave satisfactory overall appearance. We use this example to emphasize that our method is not a straight transfer of colors in the traditional sense, but rather a transfer of an "illuminant impression", which often gives compelling compositions that seem to be hand-tuned, without causing colors of skin and natural objects to be distorted in unnatural ways.

Summary

We proposed and implemented a set of color and tone transfer algorithms that enable color adjustment of consumer images under a strong naturalness constraint. Color properties of input images are estimated and described with a small set of descriptors based on plausible illuminant at the scene. Color transfer is performed as a re-adaptation process using the RLAB color appearance model. The use of a color adaptation model to perform color transfer enables us to impose a naturalness constraint on the adjustment process without explicitly defining a naturalness metric. Our method is neither a white point estimation and correction process, nor a direct transfer of target colors to source images. Instead, we estimate and transfer an illuminant "impression" of the image. This makes it possible for us to treat artwork images the same way as photographs, and use artwork as color transfer target. In the case of a single-photo composition, we can easily reverse the process and modify artwork colors to match the user photo in the same way.

References

[1] Erik Reinhard, Michael Ashikhmin, Bruce Gooch and Peter Shirley, "Color Transfer between Images", IEEE CG&A special issue on Applied Perception, Vol 21, No 5, pg. 34-41. (2001).

[2] X. Xiao and L. Ma. "Color transfer in correlated color space", Proc. ACM international conference on Virtual reality continuum and its applications, pg. 305–309. (2006).

[3] M. D. Fairchild, and R. S. Berns, "Image Color-Appearance Specification through Extension of CIELAB", Color Res. Appl. 18, pg. 178-190. (1993).

[4] Xiaodi Hou and Liqing Zhang. "Color conceptualization", Proceedings of the 15th international conference on Multimedia, ACM, New York, NY, USA, pg. 265-268. (2007)

[5] J. Vrhel, R. Gershon, and L. Iwan, "The Measurement and Analysis of Object Reflectance Spectra", Color Research and Application, vol. 19, no. 1, pg. 4-9. (1994).

[6] Peter B. Delahunt, Xuemei Zhang and David H. Brainard, "Perceptual image quality: Effects of tone characteristics", J. Electron. Imaging 14, 023003 (2005)

Author Biography

Xuemei Zhang received her BS in psychology from Beijing University (1988) and her Masters in statistics and PhD in psychology from Stanford University (1997). Since then she has worked as an imaging scientist/architect at HP, Agilent, Micron, and currently at Apple. Her work has focused on imaging algorithms for cameras and displays. This work was done while she was a senior research scientist at HP Labs.

Hui Chao is currently a Senior Staff Engineer at Qualcomm Corporate Research and Development. She received her BS in Electric Engineering from Wuhan University in China and her PhD in Medical Imaging from Massachusetts Institute of Technology. She was a senior research scientist at HP Labs while this work was carried out.

Dan Tretter received B.S. degrees in mathematics and electrical engineering from the Rose-Hulman Institute of Technology in Terre Haute, IN. He earned his M.S.E.E. and Ph.D. degrees from Purdue University. Dan is currently a Research Manager at HP Labs, where his project team develops technologies that leverage multimedia analysis for improved media management and creation of custom media presentations. Dan's current research interests include the management, browsing, and presentation of personal media collections.