Color Correction of Smartphone Photos with Prior Knowledge
Jun Jiang\textsuperscript{a}, Yonghui Zhao\textsuperscript{b} and Shen-ge Wang\textsuperscript{b}
\textsuperscript{a}Munsell Color Science Laboratory, 54 Lomb Memorial Drive, Rochester, NY14623;
\textsuperscript{b}Xerox Research Center Webster, 800 Phillips Rd, Webster, NY 14580

ABSTRACT

Human visual system has the property of perceiving the object color to remain constant regardless of the prevailing illumination. However, digital cameras usually lack this capability, and the captured images are digitally corrected to discount the color of the scene light based on the estimated illuminant. Illumination estimation might be erroneous in some artificial or chromatic lighting conditions. A method was proposed to correct digital photos captured with a smartphone camera using the smartphone owner’s face as the reference. Taking the advantage of the latest smartphones with two build-in cameras, we could use the front camera to capture the smartphone owner’s face and compare with the saved reference face image in order to estimate the scene illuminant. After that, we could properly adjust the capture setting for the main camera in order to take a decent target image; or we could automatically correct the target image based on the estimated illumination by comparing two face images. The method was implemented on the iOS mobile platform. Experimental result shows that the adjusted images using the proposed method are generally more favorable than the pictures taken directly by the default camera application.

Keywords: Color Constancy, Chromatic Adaptation, White Balance, Contrast Adjustment

1. INTRODUCTION

A color image is the result of a complex interaction between three major components: the scene, the illumination and the camera sensor. The change in any of them usually leads to a different color perception. Typically, the scene and the illumination are tightly coupled and very hard to separate from each other. Human visual system has the property of perceiving the object color to remain constant regardless of the prevailing illumination. This perceptual ability is called color constancy.\textsuperscript{1} For example, a copy paper is usually perceived by a typical observer as white despite of illuminant change from the direct sunlight in the noontime to the twilight during the sunset. But, digital cameras usually lack of this capability.\textsuperscript{2} There are some built-in features available on most cameras to automatically configure white point and exposure time; however, in some cases, the photos are still not satisfactory because of unwanted color cast or poor lightness contrast.

Figure 1. Three baby photos taken with iPhone 3GS under different illumination conditions
Figure 1 shows three baby photos taken under different illumination conditions. The left image was taken under bright sunlight, while the other two images were captured indoor under tungsten light and compact fluorescent light, respectively. The baby face is quite normal and pleasing in the left image, while the same face looks slightly yellowish or reddish in the other two images. So the build-in AWB (automatic white balance) did an imperfect job of estimating the scene illuminant and then correcting the image to discount the color of the illuminant.

A method was proposed to automatically remove unwanted color cast and to adjust poor lightness contrast of digital photos. With the increasing popularity of smartphones, it is becoming more beneficially to implement this method for smartphone cameras. Most people carry their smartphones all the time, and would like to snap pictures spontaneously whenever and wherever. Due to miscellaneous illumination conditions, some photos might be poorly color balanced and contrast adjusted. Thus, it is desirable to use a reference with known optical property to estimate a scene illuminant, and the proposed method is to use the smartphone owner’s face as the reference. Taking the advantage of the latest smartphones with two build-in cameras, we could use the front camera to capture the smartphone owner’s face and compare with the saved reference face image in order to estimate the scene illuminant. After that, we could properly adjust the capture setting (e.g., white point, exposure time, and flash on/off) for the main camera in order to take a decent target image; or we could automatically correct the target image based on the estimated illumination by comparing two face images.

2. PROPOSED METHOD

2.1 White Balance Adjustment

The AWB functionality can be found in almost every camera, from high-end single-lens reflex (SLR) camera to inexpensive webcam. The complexity of AWB algorithms and image quality after adjustment may vary a lot depending on camera manufacturers and captured scene contents. Despite of their differences, AWB algorithms are usually implemented in two ways. The first type of the algorithms is based on a diagonal model, derived from the von Kries Hypothesis\(^3\) that white balance is an independent gain regulation of three cone signals through three different gain coefficients, as shown in Eq. (1),

\[
\begin{bmatrix}
L' \\
M' \\
S'
\end{bmatrix}
= \begin{bmatrix}
k_L & 0 & 0 \\
0 & k_M & 0 \\
0 & 0 & k_S
\end{bmatrix}
\begin{bmatrix}
L \\
M \\
S
\end{bmatrix}
\tag{1}
\]

where \(L, M,\) and \(S\) are three cone signals responding at long, middle and short wavelength ranges, respectively, \([L \ M \ S]^T\) and \([L' \ M' \ S']^T\) are the cone signal vectors before and after white balance adjustment, and \(k_L, k_M\) and \(k_S\) are three gain coefficients to account for the change in the illuminant. Grey-world assumption (GWA) belongs to this group. It assumes that given an image with sufficiently varied colors, the average surface color in a scene is gray.\(^4\) GWA usually works reasonably well in natural scenes, but the result may be far from satisfaction once the assumption is invalid. The second type of the algorithms tries to estimate correlated color temperature (CCT) of a scene,\(^3,7\) which indicates the illuminant type, e.g., CIE standard illuminant D65 (CCT=6500K). These algorithms may also work well in natural scenes, but may be less effective when pictures are taken under artificial or chromatic lighting conditions.

Instead of relying on pure assumption or CCT estimation, it is desirable to use a reference with known optical property to estimate a scene illuminant. The proposed method is to use the smartphone owner’s face as the reference. A reference face image was taken in advance under bright sunlight or sufficient daylight simulation and saved in the photo library. Whenever the user snaps a target photo, he should also take an image of his own face, referred to as the current face image in the following discussion. Then, common features are extracted from both the reference and current face images and compared to build a diagonal model for white balance adjustment. In this paper, the highlight on the nose is used mainly because the specular highlight on the nose tip is illuminated by the direct light without suffering from shadows or inter-reflections from other objects, thus more likely revealing the lighting information. The highlight on the nose tip is identified by the pixel with the maximum luminance value. Within the cropped nose range, the luminance value for each pixel is calculated using the formula,

\[
Y = 0.3 \times R + 0.59 \times G + 0.11 \times B
\tag{2}
\]
where $Y$ is the calculated luminance value based on the weighted sum of pixel values for red, green and blue channels, and these weights are obtained from CCIR 601 standard. After that, a diagonal model is applied to the target photo using Eq. (3),

\[
\begin{bmatrix}
R'_t \\
G'_t \\
B'_t
\end{bmatrix} =
\begin{bmatrix}
G_{\text{ref}}/G_{\text{cur}} & 0 & 0 \\
0 & G_{\text{ref}}/G_{\text{cur}} & 0 \\
0 & 0 & G_{\text{ref}}/G_{\text{cur}}
\end{bmatrix}
\begin{bmatrix}
R_t \\
G_t \\
B_t
\end{bmatrix}
\]

(3)

where $[R_t \ G_t \ B_t]^T$ and $[R'_t \ G'_t \ B'_t]^T$ are the camera responses of the target photo before and after adjustment, and three gain coefficients in Eq. (3) are the channel-wise ratio of the pixel values between the reference and current face images. In order to maintain the luminance of the target image, the scaling factor $G_{\text{cur}}/G_{\text{ref}}$ is multiplied so that the green channel of the target photo is not altered. The pixel with the maximum luminance should not be saturated. To reduce the noise and increase the robustness, the average value of the top 2% pixels with the largest luminance should be used in Eq. (3) to calculate the gain coefficients, instead of a single pixel with the maximum luminance.

### 2.2 Lightness Contrast Adjustment

In addition to white balance adjustment, the reference can also be used to correct lightness contrast of pictures. Contrast enhancement is a process that allows some image features to show up more visibly by making the best use of the lightness range presented on the display device. Generally, contrast can be adjusted either in the spatial or in the frequency domain. The histogram equalization and sigmoid function are among the most popular methods to tune the contrast in the spatial domain.

The sigmoid function was used in the paper to create the better contrast of the human face in the picture while other objects in the scene might be slightly over-exposure or under-exposure. The sigmoid function is applied to the luminance channel only without affecting the chrominance channels, using the formula,

\[
Y'' = \frac{1}{1+e^{-c(Y-m)}}
\]

(4)

where $c$ is the desired contrast, $m$ is the inflection point around which lightness contrast is the most effectively boosted, $Y'$ and $Y''$ are the luminance values before and after adjustment, respectively. The greater the $c$ in Eq. (4), the steeper the change in the contrast near the inflection point, as shown in Figure 2.

![Figure 2. The sigmoid curves with different contrast values (m = 0.5)](image)

### 2.3 User Interface Design

The method was implemented on the iOS mobile platform. The user interface with two views is shown in Figure 3. Pictures could be taken via the capturing view through the viewfinder [Figure 3 (a)]. The top right corner has two buttons to lock the current white point and expose time, respectively. During a capture session, both the white point and exposure time could be locked to a fixed known state. However, the locked state will be changed after switching two cameras of a smartphone. So, it means that the front and main cameras could not be locked to the same capture setting. It is a limitation of camera hardware and accessible APIs. Thus, for our implementation, we only use the main camera to capture both the current face image and the target photo. On the other hand, white balance and contrast adjustment could be done via the processing view [Figure 3(b)]. First, three images including the target photo, the reference and current face images are loaded into the memory using three buttons. Then, four adjustment methods could be tested by clicking...
one of four buttons, and they are (1) lightness contrast adjustment (“Contrast” button), GWA (“GWA” button), white balance adjustment (“Prior” button) and the combination of (1) and (3) (“Go” button).

3. EXPERIMENTS

3.1 White Balance Adjustment

Figure 4(a) was the reference face image taken under daylight simulator in a light booth, while the current face image [Figure 4(b)] and the target photo [Figure 4(c)] were captured under the horizon light.

In Figure 4, the current and target images taken under the horizon light appeared much more yellowish, resulting from the higher energy at the longer wavelength of the horizon light source. White balance had to be adjusted to remove the yellowish colorcast in the current and target images.

Face detection and feature extraction were implemented using OpenCV. The nose regions from both the reference and current face images were identified and cropped out. The detected nose regions were shown in Figure 5. Saturate pixels were removed from the further calculation. The top 2% pixels with the largest luminance values were selected and averaged. After that, the gain coefficients were calculated as the ratios of the average pixel values of the reference and current face images, and then used to adjust the target image [Eq. (3)]. The corrected target image was shown in Figure 4(d).
Figure 5. The nose regions extracted from the reference (a) and current (b) face images using OpenCV.

Figure 4 (c) was taken by the default camera application on the iPhone 4 with AWB on. Despite of built-in AWB, the picture still appeared quite yellowish. Noticeable improvement could be found in Figure 4(d) adjusted with the proposed method. The paper was as white as it should appear after adjustment. Moreover, the color appearance of the 24 color patches was more satisfactory.

3.2 Lightness Contrast Adjustment

Image lightness contrast could also be adjusted with a sigmoid function defined in Eq. (4). The parameter \( c \) was empirically set to be 2.5. In order to better enhance the contrast of the face region in the picture, the inflection point \( m \) was set close to the mean luminance of the face tone. The mean luminance in the cropped nose region was used to obtain \( m \) using the formula,

\[
m = k \frac{L_{\text{cur}}}{L_{\text{ref}}}
\]

where \( L_{\text{ref}} \) and \( L_{\text{cur}} \) were the mean luminance values of the extracted nose regions in the reference and current face images, respectively, and \( k \) was a scalar factor to make minor adjustment. As the nose usually appeared lighter than other parts of the face, \( k \) was set to be 0.5 empirically. The result was shown in Figure 6.

In Figure 6, the original picture was pretty dim, resulting from under-exposure. The adjustment was made based on the reference image [Figure 4(a)] and the original image [Figure 6(a)]. In this example, the original picture was used as both the current and target images. Better lightness contrast was achieved after the adjustment, and the adjusted picture [Figure 6(b)] was much more pleasing.

4. DISCUSSIONS

4.1 Color Space

White balance adjustment was implemented in the device RGB space for fast speed, while the lightness contrast was adjusted in device-independent CIE XYZ space. The XYZ color space was found to be a better space to correct the white balance of the picture. \(^{13}\) However, the improvement of XYZ color space over device RGB space through additional color transformation between these two spaces was unnoticeable in our pilot experiment. Therefore, the device RGB space was selected in the implementation to adjust white balance for both computational efficiency and image quality reasons. On the other hand, the contrast adjustment was made in the XYZ space rather than RGB color space, mainly because
most of the contrast in natural scenes was perceived in the luminance channel rather than in the chromaticity channels, and all three channels in the RGB space contribute to the perception of lightness.

### 4.2 Camera Limitation

The quality of smartphone cameras had a big impact on the captured images in the end. It was critically important when the luminance contrast was boosted. The contrast boosting would make the darker regions appear mush noisy and even chromatic. Therefore, the noise existing in the image had to be probably handled. Otherwise, artifacts would inevitably be shown in the darker regions of the picture. Some average or smoothing operation was implemented in the dark region to make these artifacts less visible.

The current implementation was made on the iOS 4.3.3. Given the limitation of the published APIs by Apple®, a full control over the camera hardware was not possible. For example, the front and main cameras could not be locked at the same time to a fixed state (e.g., white point, exposure time, and etc.). Future development could be made on a more open-sourced platform, if more controls over the hardware are desired.

### 5. CONCLUSIONS

A new method was developed to adjust white balance and lightness contrast of the smartphone photos based on the prior face information. One decent face image was taken in advance under sufficient illumination condition and saved as the reference picture. Whenever a person snapped a target photo, he also took an image of his own face as the current face image. The reference and current face images were compared to estimate the capturing illumination and to build the proper color transformation. After that, the transformation was used to correct the target photo taken under the similar illumination condition as the current face image. It was implemented on the iOS mobile platform. Experimental result shows that the adjusted images are generally more favorable than the pictures taken directly by the default camera application.

### REFERENCES


