COLOR CORRECTION VIA ROBUST REFERENCE SELECTION AND RECOVERY USING A LOW-RANK MATRIX MODEL

Dong Li\(^1\), Xudong Xie\(^2\) and Kin-Man Lam\(^1\)

\(^1\)Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong, China
\(^2\)Department of Automation, Tsinghua University, Beijing, China
dong.li@polyu.edu.hk, xdxie@tsinghua.edu.cn, enkmlam@polyu.edu.hk

ABSTRACT

In this paper, we propose a method that can handle the color correction of a large collection of photographs simultaneously and automatically via robust reference selection. The method does not use any particular model to handle the errors on the photographs, but corrects all kinds of errors caused by changes of viewpoint, large illumination variations, gross pixel corruptions, and partial occlusions under a low-rank matrix model. Furthermore, our method uses the image pixel values directly in vector form, which preserves the spatial information, to obtain the matrix for color correction, unlike other statistics-based image-representation methods such as color histograms. Experiments verify that our method can achieve consistent and promising results on uncontrolled real photographs acquired from the Internet.

Index Terms— Color correction, color constancy, low-rank matrix model, rank minimization

1. INTRODUCTION AND MOTIVATION

Due to different illuminations of the scenes and different cameras’ properties, the same object may appear to show very different colors in different photographs. Hence, in order to recover the true color of the objects, invariant to the color of the light sources, many methods have been proposed. The color constancy methods usually use certain hypothesis in making color corrections. For instance, Grey-Edge [1] assumes that the average edge difference in a scene is achromatic. These methods only use the information from a single image, so they may sometimes enlarge the color discrepancy between the same objects. Color correction or color transfer method is the process of correcting or transforming the colors of a source image to the colors of a target image. Reinhard et al. [2] used a linear transform which changes the statistics of color distributions of two images. Kim and Pollyfeys [3] used a look-up table derived from the 2D joint histogram of image feature correspondences or pixel pairs in the overlapped area of two images. These methods use the information from both the source and the target images.

However, it is usually difficult to choose a source or a target image as a suitable reference, which will be vital to the quality of the color-corrected images. In practical applications, photographs will suffer from large illumination variations, partial occlusions, and/or gross pixel corruptions (e.g. shadows, specularity).

Although images of an object may be corrupted in different ways or suffer from different transformations, we can still make proper corrections of these images by considering several, or even hundreds, of them at the same time. This is made possible by collecting similar images from the Internet, or by taking advantage of digital cameras with a large memory capacity. For instance, when travelling, one may have taken hundreds of photos, including one’s own face. In addition, millions of photographic records of popular landmarks, such as the Statue of Liberty in New York and the Trevi Fountain in Rome, have been uploaded and made available on the Internet. The images of a particular object (face or landmark) in these photographic records can be stacked as columns of a matrix, which should be low-rank [4], contain sufficient information to recover the original data, and provide a basic measure and constraint for color correction.

Recent advances in rank minimization have shown that it is indeed possible to recover low-rank matrices efficiently and precisely, even when the data are significantly corrupted, by using tools from convex programming [4, 5]. In this paper, we consider the problem of robustly extracting color information in a low-rank model for automatic or interactive color correction. Our method need not use a particular model to handle a particular error, and can model robustly the errors caused by viewpoint changing, large illumination variations, gross pixel corruptions and partial occlusions under a low-rank matrix framework. Furthermore, by

---

* This work was supported by a grant from The Hong Kong Polytechnic University, Hong Kong, the project 60872085 of NSFC, and an Open Project Program of the National Laboratory of Pattern Recognition (NLPR).
taking the advantage of robust reference selection and recovery, the correction matrix can be acquired by directly using the image pixel values in vector form, instead of other statistics-based image representation such as color histograms. This can better retain the spatial information on images. Finally, experiments were conducted based on uncontrolled real photographs collected from the Internet; this resembled real applications.

2. COLOR CORRECTION VIA ROBUST REFERENCE SELECTION AND RECOVERY

Our method works on the CIELAB color space, which is designed to approximate human vision. To separately deal with lightness and color, an input image is first converted to the CIELAB color space. For photographs with faces, face detection [6] and face recognition [7] technology can be employed; correspondingly, for landmark photographs, SIFT features [8] are extracted and similar regions of the landmark concerned are cropped after matching [9]. As the L’ channel which represents lightness contains more information about the shape, the transformation recovery is performed on the L’ channel rather than the a* and b* channels.

2.1 Lightness and transformation recovery

Modeling viewpoint variations. Because the photos may be taken by a perspective camera from different viewpoints, the color distribution will change, as shown in Fig. 1(b). Hence, all these errors have to be recovered and corrected. Since the 3-D structure of the object of interest is unknown, we assume that the viewpoint changing and misalignment are restricted to the image plane. Thus, the viewpoint changing and misalignment can be modeled using a domain transformation operator \( \tau : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \), such that

\[
I'_L(x, y) = (I_L \circ \tau)(x, y) = I_L(\tau(x, y)),
\]

where \((x, y)\) are the pixel’s coordinates, and the image \(I'_L\) is the image \(I_L\) after the alignment operation \(\tau\) is performed on the \(L'\) channel. Then, the well-aligned (transformed) images \(I'_L = I_L \circ \tau\) \((i = 1, \ldots, n)\) can be determined and stacked in columns to form a low-rank matrix \(L'\) as follows:

\[
L' = L \circ \tau = \begin{bmatrix} \text{vec}(I'_1) & \cdots & \text{vec}(I'_n) \end{bmatrix}.
\]

Hence, the shape recovery problem of the reference regions can be modeled as the following optimization problem:

\[
\min_{L, \tau} \text{rank}(L') \quad \text{s.t.} \quad L \circ \tau = L'.
\]

2.2 Color recovery

After the domain transformation \(\tau\) has been computed using (5), we can employ \(\tau\) to transform the corresponding \(a^*\) and \(b^*\) channels, as represented in the matrix \(C\), to obtain the well-aligned image matrix \(C'\) of the corresponding channels as follows:

\[
C' = C \circ \tau.
\]

Modeling corruption. As described in Section 2.1, the true color distribution of an image may be disturbed by different corruptions, caused by large illumination variations, specularity, purple fringing, etc. Similarly, we model the corruptions as a sparse error. In the \(a^*\) and \(b^*\) channels, the image similarity measure based on matrix rank can still work. By optimizing the following problem:

\[
\min_{C', \tau} \text{rank}(C') + \gamma ||E_c||_1 \quad \text{s.t.} \quad C' = C \circ \tau,
\]

the recovered reference region \(C'\) can contain not only most of the useful color statistical information, but also remove the corruptions.

Although solving the rank function and the \(\ell^1\)-norm in (5) and (7) are usually NP-hard problems, recent developments in convex optimization have shown that, under fairly broad conditions (i.e., when the rank of the matrix to be recovered is not too high and the number of errors is not too large), they can be effectively recovered via their convex surrogates. The nuclear norm or sum of the singular values replaces \(\text{rank}(\cdot)\) and the \(\ell^1\)-norm replaces \(\ell^1\)-norm [4]. Several algorithms can solve the problem, such as the Augmented Lagrange Multiplier (ALM) method [10], the Accelerated Proximal Gradient method [11], and the Singular Value Thresholding method [12]. In this paper, the ALM method is adopted because of its efficiency and accuracy.

2.3 Color statistics and transfer

Having performed the robust selection and recovery, the color information of the reference images has become more definite, and the reference images have also been well aligned to match each other at the pixel level. This can help to guarantee that using a simple color statistics and transfer technique can probably achieve a desirable result. Therefore, for the sake of efficiency, we use the von Kries model [15] rather than other, more complicated methods. It is assumed that the transformation is linear and the transformation matrix is estimated from the converted model as follows:

\[
I_t = M \times I_s,
\]

where \(I_s, I_t = \begin{bmatrix} \text{vec}(I'_s) & \text{vec}(I'_t) \end{bmatrix}^T \in \mathbb{R}^{k \times 1}\) are the source image and the target image (the target image is recovered from the low-rank matrix model), respectively, and \(M\) is a \(3 \times 3\) diagonal matrix. Unlike those methods based on histograms or statistical information about images, our algorithm directly uses the image pixel-level information to derive the correction matrix \(M\).
3. EXPERIMENTS

We will evaluate the performance of our method in two different scenes: one is images of the same person taken from different environments; and the other is photographs of a famous building. The face or the object is selected and recovered as references for color correction.

3.1 Photographs of the same person

First, our method is evaluated using the Labeled Faces in the Wild (LFW) dataset [14] of celebrity images, collected from the Internet. The face images in the dataset exhibit significant variations in pose and facial expression, in addition to illumination and occlusion. Our aim is to use the face images of the same person as references to perform color correction on the batch of photographs. Fig. 2(a) shows some photographs of Ariel Sharon in the LFW dataset taken under different conditions and environments.

An off-the-shelf face detector is employed to obtain the initial face region in each image, as shown in Fig. 3(a). These face images show very different color distributions, and have different poses. What is more, the images also suffer from expression variations, occlusion, specularity, and shadowing. Traditional color correction methods can hardly extract the correct color information simultaneously and robustly from images with such kinds of errors. Using the low-rank matrix model, the lightness and shape recovery is performed on the \( L^* \) channel, while color recovery takes place on the \( a^* \) and \( b^* \) channels using the low-rank matrix model. Fig. 3(b) shows the recovered face images which are well aligned and preserve the correct color information. In addition, our method can remove some occlusion or corruption, as shown in the magnified images in Fig. 3(c). We use the average of all the recovered images as the target image for color correction; the results are shown in Fig. 2(b). Usually, the average is less susceptible to noise. However, if a reference can be identified manually as having the correct color, it can be selected as the target. Compared with the results based on the Grey-Edge method, as shown in Fig. 2(c), we can find that our method can correct the face color more consistently.

The angular error [1] is used as a measure of the performance for color constancy. To represent the skin color, a small skin patch (3x3 pixels) is manually selected (e.g. a forehead region without specularity), which should be close to being a Lambertian surface. Thus, the skin color is approximately defined as the average of the 3x3 pixel values, which is denoted as \( \mathbf{e}_i \) for the \( i^{th} \) image. Then, we have

\[
\text{angular error} = \cos^{-1}(\mathbf{e}_i \cdot \mathbf{t}), \quad \mathbf{t} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{e}_i, \quad (9)
\]

where \( \mathbf{e}_i \) represents the normalized vector of \( \mathbf{e}_i \). Table 1 shows the mean angular errors of skin colors in the original images and in the corrected images, using our method and the Grey-Edge method based on the LFW database. The results show that our proposed method significantly outperforms the Grey-Edge method. In other words, the colors in the respective corrected images based on our method show more consistency.

| Table 1. Angular error (degrees) based on the LFW dataset |
|-----------------|-----------------|-----------------|
|                | Proposed method | Grey-Edge       | Original       |
| Median          | 2.54            | 5.2             | 4.72           |
| Average         | 2.85            | 5.6             | 5.34           |

3.2 Photographs of the same object

A colorful relief pattern was used in this part of the experiment. Entering the search term “Hall of maps Rome” on picasaweb.google.com will return more than 300 photographs of this popular landmark. Fig. 4(a) shows some examples we acquired from the website. We choose one angel pattern on the hall as the reference, which exhibits in several different colors. Before the recovery, the color distributions of the patterns are significantly different due to the variation in viewpoints, as can be seen in Fig. 5(a). All the reference regions are averaged to form the reference for the target images, and Fig. 5(b) shows the recovered images via the low-rank matrix model. As a mix of illuminations may occur in the scene, some results in Fig. 4(b) shows artifacts in bright regions. Nevertheless, the color of the hall is transformed to a uniform color.
more consistently and correctly. We compare our method with the Grey-edge method, and the results are shown in Fig. 4(c). The Grey-Edge method renders the hall with colors more randomly. The angular error measure of the methods is shown in Table 2.

![Fig. 4. Color correction based on photographs from the Picasa website.](image)

<table>
<thead>
<tr>
<th></th>
<th>Proposed method</th>
<th>Grey-Edge</th>
<th>Original</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>2.72</td>
<td>8.75</td>
<td>8.71</td>
</tr>
<tr>
<td>Average</td>
<td>2.87</td>
<td>9.76</td>
<td>9.31</td>
</tr>
</tbody>
</table>

### 4. DISCUSSION AND FUTURE WORK

In this paper, we have proposed a method which can handle a large collection of photographs for color correction simultaneously and automatically. The method does not employ a particular model to handle a particular error, but models the different kinds of errors caused by changing viewpoint, large illumination variations, gross pixel corruptions, and/or partial occlusions under the low-rank matrix framework robustly. Good experimental results have demonstrated the effectiveness of our method. However, the real world is filled with complex geometry and non-uniform illumination. As our method is based on the von Kries model, some conditions, such as mixed illuminations in Fig. 6(a) or specular reflection on the human face, will result in artifacts. Hence, to further improve the performance, a more sophisticated light-reflectance model may be employed. Hsu et al. [13] proposed a method which is able to handle a mix of two illuminants by assuming a prior knowledge about the state of the illuminants and the scene. One possible future work will be the handling of mixed illumination conditions. Also, the error identified in the recovery process can be used to guide the color correction.

### 5. REFERENCES