Computer Vision and Image Understanding 146 (2016) 77–85

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Separation of reflection components by sparse non-negative matrix factorization

Yasushi Akashi [a], Takayuki Okatani [a,b,*]

[a] Tohoku University, 6-6-01 Aramaki Aza Aoba, Aoba-ku 980-8579 Sendai, Japan
[b] JST CREST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

ARTICLE INFO

Article history:
Received 23 March 2015
Accepted 2 September 2015
Available online 14 September 2015

Keywords:
Reflection components separation
Specular reflection
Diffuse reflection
Dichromatic reflection model
Non-negative matrix factorization
NMF

ABSTRACT

This paper presents a novel method for separating reflection components in a single image based on the dichromatic reflection model. Our method is based on a modified version of sparse non-negative matrix factorization (NMF). It simultaneously performs the estimation of diffuse colors and the separation of reflection components through optimization. Our method does not use a spatial prior such as smoothness of colors on the object surface, which is in contrast with recent methods attempting to use such priors to improve separation accuracy. Experimental results show that as compared with these recent methods that use priors, our method is more accurate and robust. For example, it can better deal with difficult cases such as the case where a diffuse color is close to the illumination color.

1. Introduction

This paper considers the problem of separating reflection components (i.e., specular and diffuse reflections) in a single image. It is useful for several purposes. One is the use with photometric methods, such as shape-from-shading [1,2] and photometric stereo [3]. As these methods often assume the surfaces of objects to be perfectly diffuse, it is necessary to eliminate specular component before applying them to real objects having specular reflectance. The separation of reflection components is also useful for the visual recognition of materials of objects; the highlights extracted from images are used as features for the recognition.

A large number of studies have been conducted to develop a method for accurately and robustly separating reflection components in a single image [4–13]. Most of them assume the dichromatic reflection model, which states that the light reflected on an object surface is given by a linear sum of a specular component and a diffuse component [4]. Specifically, the 3-vector $\mathbf{i}_p$ containing the RGB values of a pixel p is given by

$$\mathbf{i}_p = \alpha_p \mathbf{i}_s + \beta_p \mathbf{i}_d, \tag{1}$$

where $\mathbf{i}_s$ is the color of the only illumination existing in the scene and $\mathbf{i}_d$ is the diffuse color (i.e., the color caused by diffuse reflection) of the object surface.

If multiple pixels p share the same illumination color $\mathbf{i}_s$ and the same diffuse color $\mathbf{i}_d$, then Eq. (1) gives constraints on the variables on the right hand side. This is the principle on which color-based methods for separating components in a single image rely. More specifically, they commonly consider the following setting:

- The object surface consists of multiple regions with different diffuse colors, each of which consists of a number of pixels with a single color $\mathbf{i}_d$.
- The illumination color $\mathbf{i}_s$ is known. The diffuse color $\mathbf{i}_d$ of each region and also which region each pixel belongs to are unknown. The coefficients α_p and β_p are different for each pixel p, both of which are also unknown.

Early studies [7,8,14] attempt to solve the problem within this setting. More recent studies [9–13,15] attempt to utilize spatial information to improve separation accuracy. To do so, they incorporate spatial priors such as the smoothness of the diffuse colors and/or the specular reflections on the object surface.

Our method separates reflection components based on sparse non-negative matrix factorization. It simultaneously performs the estimation of diffuse colors and the separation of reflection components through optimization. It is notable that our method does not use an additional prior or assumption as those used in the recent studies. In this respect, our study runs counter to the recent trend of research, which is also the argument we make in this paper. That is, the above setting with the dichromatic model (1) alone might be more sufficient than expected for accurate separation of reflection components. In fact, as shown in the experimental results, our

* Corresponding author.
E-mail address: okatani@vision.is.tohoku.ac.jp, okatani@gmail.com (T. Okatani).

http://dx.doi.org/10.1016/j.cviu.2015.09.001
1077-3142/

method is more accurate and robust than the state-of-the-art that uses additional priors. As an additional prior is not necessary, our method is free from tuning a number of hyper parameters.

2. Related work

The early approach to the problem is to determine diffuse colors by analyzing the color space. A number of studies [7,8,14] are fallen in this category, and they solve the problem in two steps: (i) they determine the diffuse colors first by analyzing the color space onto which all the image pixels are projected, and (ii) then determine the other unknowns using the results. The method of Klinker et al. [7] performs clustering of all the pixels in the RGB color space to determine diffuse colors. Bajcsy et al. [8] used the Hue-Saturation-Lightness color space instead of the RGB space. The method of Tan and Ikeuchi [14] projects pixel colors along the direction of the illumination color to a point of the lowest observed intensity to determine the diffuse color. However, all these methods tend to be vulnerable to the clutters in the color space, such as image noises and color blending along the border of diffuse colors.

To cope with this difficulty, more recent methods attempt to utilize spatial information in the image [9–13,15]. Instead of determining diffuse colors first, most of these methods search for all the parameter values simultaneously through optimization. Some of them use a specular-free image (or its extension), an image free from specular components but with distorted diffuse components. It is created from the input image usually by a simple, pixel-wise operation.

Tan and Ikeuchi [9] first proposed a method of this category. They showed a method of creating the specular-free image by setting the maximum chromaticity of each pixel to an arbitrary value. Based on this, they presented a method that iteratively separates the reflection components by using a relation of two neighboring pixels. In their method, diffuse colors are estimated gradually in such a way that information propagates from outside highlight regions to inside them. This propagation often fails on the boundary of diffuse colors. It also cannot correctly deal with diffuse colors having the same hue but different saturation. To solve these problems, Tan et al. [10] proposed a method for recovering diffuse components by using the texture information around highlights, but it requires the positions of highlights to be known.

There are a few studies [12,15] extending the method of Tan and Ikeuchi, which incorporate an explicit prior that the diffuse colors should be mostly smooth on the object surface, except for occasional region boundaries. Under this assumption, these studies propose methods that apply smoothing to the image of an only color channel that contains the specular component, smoothing it out and then separating the two components. The key issue, then, is how to prevent smoothing to be applied across different regions of different diffuse colors. Yang et al. [12] employ bilateral filtering whose range filter is determined by approximate diffuse color and apply it to the image of maximum chromaticity. Mallick et al. [15] choose anisotropic erosion whose structuring set is determined by surface texture and apply it to the S channel of the SUV color space.

There are more studies that follow a similar approach. Shen and Cai [11] proposed another specular-free image that is obtained by subtracting the minimum of the RGB values from them for each pixel. They proposed a simple separation method based on it and also on an incorporated prior that the diffuse color changes smoothly around highlights. Although it is simple and fast, their method is less accurate than the above methods, as it simplifies the problem too much, resulting in that Eq. (1) will no longer be satisfied. Kim et al. [13] have recently proposed an optimization-based approach that uses three different priors (i.e., the spatial smoothness of specular reflections and diffuse colors, and the number of diffuse colors being as small as possible). They also propose to apply the dark channel prior [16] to obtaining another specular-free image, although it is exactly the same as the one of Shen and Cai [11]. Their method alternately performs the following steps in an iterative manner: (i) cluster image pixels based on the latest estimate of their chromaticity and (ii) apply an edge-preserving filter to the result, followed by reassignment of labels. However, it remains unclear how accurate their method is, since their experiments compare mostly with the method of Tan and Ikeuchi [9] alone and not with that of Yang et al. [12], which is more close to their method in that the smoothness of diffuse colors is assumed and an edge-preserving filter is used. Moreover, their method requires a number of hyperparameters (and it is unclear how to choose them) and also the assumed three priors are too much and could narrow the range of applicability.

3. Non-negative matrix factorization (NMF)

Our method is based on sparse non-negative matrix factorization (sparse NMF). Before describing our method, this section briefly summarizes sparse NMF and its numerical algorithm.

3.1. Basic NMF

NMF is a general-purpose method for multi-variate analysis. For data consisting of non-negative values such as images and speech signals, it factorizes the data into additive components. To be specific, a $M \times N$ matrix $\mathbf{V}$ containing only non-negative elements is factored into a product of a $M \times R$ matrix $\mathbf{W}$ and a $R \times N$ matrix $\mathbf{H}$, both of which similarly contain only non-negative elements:

$$\mathbf{V} \simeq \mathbf{WH}, \tag{2}$$

or the jth column vector $\mathbf{v}_j$ of $\mathbf{V}$ is represented by a linear combination of the column vectors $\mathbf{w}_k$'s of $\mathbf{W}$ weighted by the (k, j) element $H_{k,j}$ of $\mathbf{H}$ as

$$\mathbf{v}_j \simeq \sum_{k=1}^{R} \mathbf{w}_k H_{k,j}. \tag{3}$$

The factorization is obtained by minimizing some cost $D(\mathbf{W}, \mathbf{H})$ measuring the difference between $\mathbf{V}$ and its reproduction $\mathbf{WH}$. For $D(\mathbf{W}, \mathbf{H})$, L_2 norm

$$D(\mathbf{W}, \mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_2^2 \tag{4}$$

is widely used for general purposes, so is in our method. The generalized KL divergence [17] and Itakura–Saito divergence [18] are sometimes used depending on problems.

Since an efficient iterative algorithm was developed by Lee and Seung [17], NMF has been applied to all sorts of problems, and various extensions have been made to the cost function depending on problems [18–24].

3.2. Sparse NMF

An important extension is the sparse NMF that incorporates the sparse regularization into the minimization [19]. It minimizes the following cost for the purpose of obtaining $\mathbf{H}$ having as small a number of non-negative elements as possible.

$$F(\mathbf{W}, \mathbf{H}) = \frac{1}{2}\|\mathbf{V} - \mathbf{WH}\|_2^2 + \lambda \sum_{i,j} H_{i,j}. \tag{5}$$

The second term on the right hand side follows the same relaxation as sparse coding [25] that L_0 norm is replaced by L_1 norm. The minimization of this cost results in that each data vector $\mathbf{v}_j$ is represented by a linear combination of as small a number of bases (i.e., the column vector of $\mathbf{W}$) as possible, as in sparse coding. Its difference from sparse coding is that the resulting quantities are all non-negative.

A numerical algorithm for this sparse NMF, i.e., minimizing this cost under the constraints that $\mathbf{W}$ and $\mathbf{H}$ both have only non-negative