

Multispectral Image Intrinsic Decomposition via Subspace Constraint

Qian Huang¹

cianhwang@gmail.com

Weixin Zhu¹

njuzhwx@163.com

Yang Zhao¹

zhaoyang@smail.nju.edu.cn

Linsen Chen¹

njucls@163.com

Yao Wang²

yw523@nyu.edu

Tao Yue¹

yuetao@nju.edu.cn

Xun Cao¹

caoxun@nju.edu.cn

¹School of Electronic Science and Engineering, Nanjing University²Department of Electrical and Computer Engineering, New York University

Abstract

Multispectral images contain many clues of surface characteristics of the objects, thus can be used in many computer vision tasks, e.g., recolorization and segmentation. However, due to the complex geometry structure of natural scenes, the spectra curves of the same surface can look very different under different illuminations and from different angles. In this paper, a new Multispectral Image Intrinsic Decomposition model (MIID) is presented to decompose the shading and reflectance from a single multispectral image. We extend the Retinex model, which is proposed for RGB image intrinsic decomposition, for multispectral domain. Based on this, a subspace constraint is introduced to both the shading and reflectance spectral space to reduce the ill-posedness of the problem and make the problem solvable. A dataset of 22 scenes is given with the ground truth of shadings and reflectance to facilitate objective evaluations. The experiments demonstrate the effectiveness of the proposed method.

1. Introduction

The observed spectrum of a single pixel is determined by illumination, reflectance and geometry. Shading image contains illumination condition and geometry information, while reflectance image contains the material reflectance property, which is invariant to light condition and shadow effect. The decomposition problem has been a long standing problem in both computer graphics and computer vision applications. For instance, shape-from-shading algorithms could benefit from an image with only shading effects, while image segmentation would be easier in a world without cast shadows.

Obviously, intrinsic image decomposition is an ill-posed problem, since there are more unknowns than observations.

In order to solve this problem, many works [30, 31, 12] focus on sparse representation spatially, but this does not hold for images in general. This paper addressed the problem of the recovery of reflectance and shading from a single multispectral image, namely, the Intrinsic Image Decomposition problem of a whole multispectral image captured under general spectral illumination, hereafter referred to as the IID problem. This problem is worth exploring since geometry and reflectance information are useful under certain circumstances, but one of them always interferes the detection of the other one. Unfortunately, growing dimensions of data make this problem harder to cope with.

The subspace constraint that we propose is based on the assumption that the reflectance and shading vectors both live in a low dimensional subspace along the spectral domain. According to the inherent nature of the multispectral image, we derive shading basis on the knowledge of illumination condition and derive the reflectance subspace basis by means of principle component analysis (PCA). Assuming that the Retinex theory [23] would continue to take effect in multispectral domain, we propose a subspace-based model so that deriving reflectance and shading from a multispectral image can be modelled as an convex optimization problem. In a significant departure from the conventional approaches which operate in the logarithmic domain, we directly operate on the image domain. The flowchart of our proposed MIID algorithm is shown in Fig. 1.

To overcome the lack of ground truth data of shading and reflectance, we provide a ground-truth dataset for multispectral intrinsic images to enable quantitative evaluation of various intrinsic decomposition methods. Quantitative and qualitative experiments on our dataset have demonstrated that the performance of the MIID method is better than prior work for multispectral images. Our work can bring merits to multiple applications, such as recolorization, relighting, scene reconstruction and image segmentation.

Our major contribution can be summarized as follows:

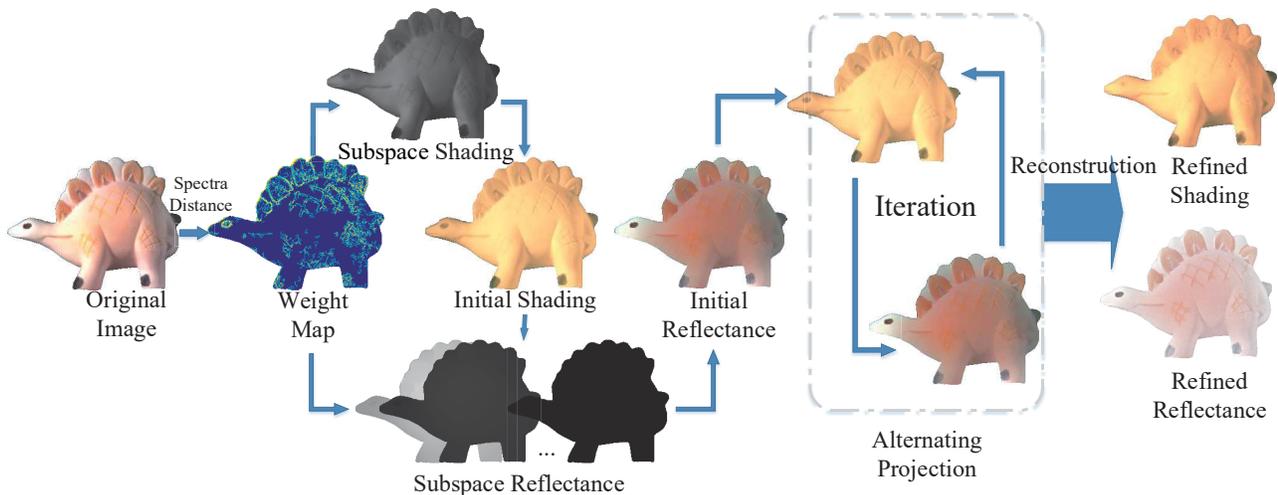


Figure 1. The flowchart of proposed MIID algorithm. Note that multispectral images in this and other figures are rendered in RGB color.

(1) we extend the Retinex model to multispectral image intrinsic decomposition, and propose a subspace constraint to handle the ill-posedness of the problem; (2) we provide a ground-truth dataset that contains the ground truth shading and reflectance images associated with captured multispectral images, which can facilitate future evaluation and comparison of multispectral image intrinsic decomposition algorithms; (3) the proposed method achieves promising results on various images both quantitatively and qualitatively.

2. Related Work

RGB Intrinsic Image Decomposition. The problem of Intrinsic Image Decomposition (IID) was first introduced by Barrow *et al.* [4]. The reflectance describes illumination-invariant albedo of the surface, while the shading contains surface geometric and illuminative information.

Some prior methods take advantage of additional information, including images sequences [34, 22, 24] to avoid shadow effect in poor lighting condition. With the improvement of sensing devices like kinect, depth cue [2, 11, 24] or surface normal [28] have been applied to infer the shading. More recently, Bousseau *et al.* [8] proposed a user-assisted method to further improve the result of separation.

Many prior methods have been proposed for the separation task from a single color image. Bell *et al.* [6] developed a dense conditional random field (CRF) based intrinsic image algorithm for images in the wild. Barron *et al.* [3] introduced shape, illumination and reflectance from shading model which performs well on images of segmented objects. Sai *et al.* [7] proposed L1 Image Transform model for scene-level intrinsic decomposition. Entropy method [13] raised by Finlayson *et al.* offered us a new viewpoint to understand this problem. With the abundance and availability of datasets and the development of computational equipment, training-based models [5, 33, 32, 38] have been built to derive reflectance and shading from RGB color images.

An especially well-known and widely employed model called Retinex [23] makes an assumption that when there is large change in chrominance, shading is usually constant, and vice versa. With Retinex theory, we are able to pinpoint where the reflectance changes in local area. Horn *et al.* [16] analyzed local derivatives for distinguishing between image variations that are due to shading or reflectance.

To overcome the drawbacks of ambiguity in local analysis, lots of research have been done to reduce the ambiguity of both reflectance and shading. Shen *et al.* [30] proposed a global optimization algorithm which combines Retinex theory and non-texture constraint to obtain global consistency of image structures. Shen [31] further applied sparse representation of reflectance as global constraint of their observation. Material cues [27] has also been introduced. But these works are only confined to RGB color images.

Trials in Multispectral Domain. Researchers have also extended trichromatic color constancy models to decompose multispectral images. A lot of trials have been made to explore this area. For example, Ikari *et al.* [18] showed us the possibility of separating illumination and surface spectral from multiple color signals. Huynh *et al.* [17] assumed that the scene could be segmented into several homogeneous surface patches, and were able to estimate the illumination and reflectance spectra under the dichromatic reflectance model. In remote sensing area, Kang *et al.* [20] fit multispectral data into trichromatic model to extract features. These works bring us new thoughts about intrinsic decomposition problems.

As for multispectral image intrinsic decomposition, Chen *et al.* [12] used super-pixel to cut down the number of unknown parameters in this underdetermined problem. Unlike the approaches above, we assume that both shading and reflectance spectral vectors live in low dimensional subspace. The subspace of shading is widely acknowledged and exploited in prior work [16], and the low-dimensional

subspace model of reflectance is introduced by [26, 29, 37]. With ground truth reflectance and illumination spectra, we can derive their subspace basis vectors and solve the subspace approximations of the true reflectance and shading images effectively.

Dataset. To establish ground-truth for intrinsic images, Tappen *et al.* [33] created small sets of both computer-generated and real intrinsic images with three color components. The computer-generated images consisted of shaded ellipsoids with piecewise-constant reflectance. The real images were created using green marker on crumpled papers [32]. Grosse *et al.* [15] provided a large dataset with three color components which is widely used in following analysis. Bell *et al.* [6] also introduced Intrinsic Images in the Wild, a large scale, public dataset of color images for evaluating intrinsic image decompositions of indoor scenes. Going beyond the three color components, [35, 10] provided a set of multispectral images of various objects without ground truth of reflectance images, and Chen *et al.* [12] built a dataset with low spectral resolution and limited diversity in image content. To the best of our knowledge, there are no other public multispectral image datasets with ground truth for multispectral image intrinsic decomposition.

3. Our Model

We assume the object surface as Lambertian and hence has diffuse reflection. In most prior work on intrinsic image decomposition, the captured luminance spectrum at every point l_p is modelled as the element-wise product of Lambertian reflectance spectrum r_p and shading spectrum s_p , where s_p is used to characterize the combined effect of object geometry, illumination, occlusion and shadowing. Mathematically, this model can be expressed as

$$l_p = s_p \cdot * r_p \quad (1)$$

where l_p , r_p and s_p are all vectors with dimensions equal to the number of spectral bands, K , of the captured image, $*$ denotes element-wise multiplication. The problem is to derive s_p and r_p from observed multispectral luminance vector l_p . At first, we will focus on recovering the reflectance spectrum using this model. Once r_p is determined, the shading image can be derived by point-wise division.

Different from the conventional approaches which operate in the logarithmic domain, we directly formulate the problem in the image domain, and this can overcome numerical problems caused by the logarithmic transformation of the image values, where noise in pixels with low intensity values can lead to large variations. Besides, although there have been substantial evidence of the subspace of the reflectance, it is not clear whether the logarithmically transformed reflectance still lives in a subspace. This makes it

hard to incorporate the subspace prior in formulations based on log-transformed images.

3.1. Estimate Reflectance or Shading Independently

The Retinex model makes following two important observations:

- 1) When there is significant reflectance change between two adjacent pixels p and q , the shading is typically constant. This leads to the relation $l_p \cdot / l_q = r_p \cdot / r_q$, where $\cdot /$ denotes element-wise division;
- 2) When the expected reflectance difference between two pixels is small, the recovered reflectance difference between the two pixels should be small.

Noting that the ratio relationship in observation 1), the ratio relationship can be written as $l_p \cdot * r_q = l_q \cdot * r_p$, or $L_p r_q = L_q r_p$ where L_p is a diagonal matrix consisting of spectral elements in l_p , we can formulate the problem of recovering the reflectance image as minimizing a weighted sum of the two energy functions:

$$E_{\text{refl}} = \sum_{p,q \in \mathcal{N}_{sc}} \|w_{p,q}(L_p r_q - L_q r_p)\|_d^d + \lambda \sum_{p,q \in \mathcal{N}_{rc}} \|v_{p,q}(r_p - r_q)\|_d^d \quad (2)$$

where \mathcal{N}_{sc} denotes shading neighborhood pair sets, \mathcal{N}_{rc} denotes reflectance neighborhood pair sets and $w_{p,q}$ and $v_{p,q}$ denote weights. $w_{p,q}$ should be large but $v_{p,q}$ be small when the expected reflectance difference between two adjacent pixels p and q is large, and vice versa. To make the formulation general, we use d to indicate the error norm.

If we directly solve for r_p , the above energy function can be written as the sum of K terms, one for each spectral component and each term can be separately minimized. With a little exercise, it can be shown that the minimal is achieved exactly when $r_p = l_p$; This is due to the inherent ambiguity of the problem, when no other constraints are imposed on r_p . We reduce the ambiguity by exploiting the fact that the reflectance spectra of typical object surfaces live in a low dimensional subspace of R^K , so that any reflectance vector can be written as a linear combination of J_r basis, with $J_r < K$.

Let B_r represent the $K \times J_r$ basis matrix for representing the reflectance vector, r_p can be written as $r_p = B_r \tilde{r}_p$. The energy function in Eq.(2) now becomes:

$$E_{\text{refl}} = \sum_{p,q \in \mathcal{N}_{sc}} \|w_{p,q}(L_p B_r \tilde{r}_q - L_q B_r \tilde{r}_p)\|_d^d + \lambda \sum_{p,q \in \mathcal{N}_{rc}} \|v_{p,q}(B_r \tilde{r}_p - B_r \tilde{r}_q)\|_d^d \quad (3)$$

The combined energy can be represented in a matrix form as:

$$E_{\text{refl}} = \|W_{L,B_r} \tilde{R}\|_d^d + \lambda \|V_{B_r} \tilde{R}\|_d^d \quad (4)$$

where \tilde{R} consists of \tilde{r}_p for all pixels in a vector. The matrix W_{L,B_r} depends on the neighborhood \mathcal{N}_{sc} considered,

the weight $w_{p,q}$, the reflectance basis B_r used, and importantly the luminance data l_p ; whereas the matrix V_{B_r} depends on the neighborhood \mathcal{N}_{rc} considered, the weight $v_{p,q}$ and the reflectance basis B_r used. Therefore, with this non-logarithmic formulation, we encode the constraint due to the measured luminance data in the matrix W_{L,B_r} .

The ambiguity with the scaling factor is inherent in all intrinsic image decomposition problems since only the product of reflectance and shading is known. To circumvent the ambiguity about the scaling factor, we further assume that the reflectance have small deviation from the input image and add a scale constraint along the entire image as $M_{B_r}\tilde{R} = C_r$, and augment the original energy function to enforce this constraint:

$$E_{\text{refl}} = \|W_{L,B_r}\tilde{R}\|_d^d + \lambda_1 \|V_{B_r}\tilde{R}\|_d^d + \lambda_2 \|M_{B_r}\tilde{R} - C_r\|_d^d \quad (5)$$

where M_{B_r} is a block-diagonal matrix depending on B_r , and C_r is a long vector consisting of luminance coefficient vectors of all l_p .

Similarly, the subspace of the shading is also widely acknowledged and exploited in prior work [14]. Shading inherently lives in a subspace, because there are usually only a few lighting sources with different illumination spectra acting in each captured scene, and the shading effect due to geometry and shadowing only modifies the spectra by a location-dependent scalar. If there is a single illumination source and its spectrum is known or is able to be identified by method of [37], we will use this spectrum (after normalization) as the only shading basis vector ($J_s = 1$ and B_s equals to this normalized spectrum). Likewise the problem of recovering the shading image can be formulate as minimizing

$$E_{\text{shad}} = \|W_{B_s}\tilde{S}\|_d^d + \lambda_1 \|V_{L,B_s}\tilde{S}\|_d^d + \lambda_2 \|M_{B_s}\tilde{S} - C_s\|_d^d \quad (6)$$

3.2. Simultaneous Recovery

Based on the formulation that solves reflectance and shading respectively, we propose an optimization algorithm that simultaneously solves both shading and reflectance. We assume that the subspace of the shading and reflectance are known, represented by basis matrices B_s and B_r , respectively, so that $s_p = B_s\tilde{s}_p$ and $r_p = B_r\tilde{r}_p$. We will use \tilde{S} to denote the long vector consisting of shading coefficient vectors \tilde{s}_p at all pixels, and \tilde{R} to denote the long vector consisting of reflectance coefficient vectors \tilde{r}_p . We propose to solve \tilde{s}_p and \tilde{r}_p , or equivalently \tilde{S} and \tilde{R} , by minimizing a weighted average of the following energy terms.

When shading is expected to be similar in pixels p and q , we have $s_p \approx s_q$ and $l_p * r_q \approx l_q * r_p$, or $L_p r_q = L_q r_p$, where L_p is a diagonal matrix consisting of spectral

elements in l_p . We formulate the energy functions directly:

$$E_{\text{sc}} = \sum_{p,q \in \mathcal{N}_{sc}} \left(\|w_{p,q}(L_p r_q - L_q r_p)\|_d^d + \|w_{p,q}(s_p - s_q)\|_d^d \right) \\ = \|W_{L,B_r}\tilde{R}\|_d^d + \|W_{B_s}\tilde{S}\|_d^d \quad (7)$$

When reflectance is expected to be similar in pixels p and q , we have $r_p \approx r_q$ and $l_p * s_q \approx l_q * s_p$, leading to a regularization energy

$$E_{\text{rc}} = \sum_{p,q \in \mathcal{N}_{rc}} \left(\|v_{p,q}(L_p s_q - L_q s_p)\|_d^d + \|v_{p,q}(r_p - r_q)\|_d^d \right) \\ = \|V_{L,B_s}\tilde{S}\|_d^d + \|V_{B_r}\tilde{R}\|_d^d \quad (8)$$

The inherent data constraint $l_p = s_p * r_p$ leads to another energy function:

$$E_{\text{data}} = \sum_p \|s_p * r_p - l_p\|_d^d = \|Q_{\tilde{S}}\tilde{R} - L\|_d^d \\ = \|Q_{\tilde{R}}\tilde{S} - L\|_d^d \quad (9)$$

where $Q_{\tilde{S}}$ is a block diagonal matrix that depends on the solution for \tilde{S} and the basis matrices B_s and B_r (likewise $Q_{\tilde{R}}$), and L is a diagonal matrix consisting of spectral elements of all input pixels.

The problem is to find \tilde{S} and \tilde{R} that minimizes a weighted average of the three energy functions:

$$E = \|W_{L,B_r}\tilde{R}\|_d^d + \|W_{B_s}\tilde{S}\|_d^d + \lambda_1 \left(\|V_{L,B_s}\tilde{S}\|_d^d + \|V_{B_r}\tilde{R}\|_d^d \right) \\ + \lambda_{\text{data},1} \|Q_{\tilde{S}}\tilde{R} - L\|_d^d + \lambda_{\text{data},2} \|Q_{\tilde{R}}\tilde{S} - L\|_d^d \quad (10)$$

Direct solution of the above problem solving \tilde{R} and \tilde{S} simultaneously is hard because of the bilinear nature of the data term. We apply the iterative solution, where we solve \tilde{R} and \tilde{S} using alternating projection. As the dimension of the shading subspace is likely to be smaller than the dimension of the reflectance subspace, we solve the shading \tilde{S} first. Also there are typically more subregions in an image with similar reflectance, where it is easier to use the constant reflectance constraint to resolve the ambiguity about shading.

A challenging issue is how to provide initial estimation of \tilde{S} and \tilde{R} in order to effectuate the data constraint in Eq.(9). We first obtain an initial estimate of \tilde{S} by minimizing Eq.(6), based on an assumed B_s . We then add a data constraint to Eq.(5) and determine \tilde{R} by minimizing:

$$E_{\text{refl}} = \|W_{L,B_r}\tilde{R}\|_d^d + \lambda_1 \|V_{B_r}\tilde{R}\|_d^d + \\ \lambda_2 \|M_{B_r}\tilde{R} - C_r\|_d^d + \lambda_{\text{data},1} \|Q_{\tilde{S}}\tilde{R} - L\|_d^d \quad (11)$$

to satisfy the data constraint. Finally, \tilde{S} and \tilde{R} are estimated iteratively by minimizing Eq.(10).

More specifically, the whole recovery algorithm is summarized in Algorithm 1:

Algorithm 1: MIID algorithm

- 1 **Step 1:** Get B_s from the ground-truth illumination and B_r from Munsell colors matt [25].
 - 2 **Step 2:** Assign constant-shading weights $w_{p,q}$ and constant-reflectance weights $v_{p,q}$ based on the observed multispectral images.
 - 3 **Step 3:** Solve an initial subspace estimate of the shading vector \tilde{S} by minimizing Eq.(6).
 - 4 **Step 4:** Solve an initial subspace reflectance vector \tilde{R} by minimizing Eq.(11).
 - 5 **repeat**
 - 6 **Step 5:** Solve \tilde{S} by minimizing the energy function in Eq.(10) with $\lambda_{data,1} = 0$, and using previously solved \tilde{R} .
 - 7 **Step 6:** Solve \tilde{R} by minimizing the energy function in Eq.(10) with $\lambda_{data,2} = 0$, and using previously solved \tilde{S} .
 - 8 **until** maximum number of iterations (100) is reached or the change in the energy function in Eq.(10) is below a threshold (0.01);
 - 9 **Step 7:** Reconstruct S and R .
-

4. Details

4.1. Weight Choice

Images suffering from poor light condition may contain shadow area, which would in turn bring in unnecessary edges that confuse the algorithm. Various methods are used to determined weights $w_{p,q}$ and $v_{p,q}$, including pixel gradient [14, 21, 23], hue [36], correlation between vectors [19] and learning [33]. To set the weight, we assume that the reflectance of two pixels are similar if their normalized luminances are similar, which can be measured by the correlation of the normalized luminances. Therefore, we proposed a illumination-robust and compute-friendly distance – normalized cosine distance, to measure the differences between spectra of two pixels in one neighborhood. This distance can be formulated as

$$d_{p,q \in \mathcal{N}_{sc}} = 1 - \frac{l'_p l_q}{|l_p| \cdot |l_q|} \quad (12)$$

$d_{p,q}$ is 0 when pixel p and q have same spectra, and increases when spectra of p and q are different. In order to derive weight $w_{p,q}$, we need to further magnify the difference between homogeneous and heterogeneous pixels and make it more robust to the noise. In our implementation

$$w_{p,q} = \frac{1}{1+e^{-\alpha(d_{p,q}-\beta)}} \quad v_{p,q} = 1 - w_{p,q} \quad (13)$$

α and β are parameters of sigmoid function. To set α and β , we sample values of α within [1000, 10000] and values

of β within $[10^{-5}, 10^{-2}]$ and choose the combination that performs best.

Fig. 2 shows different separation results for different beta values. If β is too small, shading tends to be more blurred; when β is too big, reflectance would be blurred.

4.2. Subspace Bases for the Shading and Reflectance Spectra

An important step in our problem formulation is to derive subspace basis matrices for the shading and reflectance spectra, respectively. With the help of multispectral imaging systems as PMIS [9] and CASSI [1], we can successfully get ground-truth illumination spectra. Also, there are plenty of works referring to how to extract illumination from images. For example, [37] can be applied in multispectral domain and performs well in implementation. In order not to complicate our method, we assume that the shading component is dominated by a single illumination source and we directly determine the shading subspace matrix B_s by using the ground-truth illumination.

For reflectance, the authors of [26, 29] have found J_r to be around 8 so as to reach the best trade-off between expression power and noise resistance in the process of fitting reflectance spectra. We set J_r to be 8 and perform Principle Component Analysis (PCA) to derive B_r from the Munsell colors matt measured by Hiltunen [25], which composes the reflectance spectra of 1269 matt Munsell color chips.

4.3. Initial Estimation

We use L2 norm for all terms, so that solving Eq.(6) or Eq.(11) is a quadratic programming problem, and can be solved efficiently using conjugate gradient method. In Algorithm 1, The solution to the unconstrained optimization problem in Step 3 satisfies the following linear equation:

$$\begin{aligned} H_s &= W_{B_s}^T W_{B_s} + \lambda_1 V_{L,B_s}^T V_{L,B_s} + \lambda_2 M_{B_s}^T M_{B_s} \\ H_s \tilde{S} &= \lambda_2 M_{B_s}^T C_s \end{aligned} \quad (14)$$

In Step 4, the linear equation can be written as

$$\begin{aligned} H_r &= W_{L,B_r}^T W_{L,B_r} + \lambda_1 V_{B_r}^T V_{B_r} + \lambda_2 M_{B_r}^T M_{B_r} + \lambda_{data,1} Q_{\tilde{S}}^T Q_{\tilde{S}} \\ H_r \tilde{R} &= \lambda_{data,1} Q_{\tilde{S}}^T L + \lambda_2 M_{B_r}^T C \end{aligned} \quad (15)$$

Because the matrix H_s and H_r is self-adjoint and sparse, we can solve this equation iteratively, which typically converges very fast. Once the parameters are preoptimized in a given data category, the estimation performance stays stable. On our dataset, we set $\lambda_1 = 100$, $\lambda_2 = 0.1$ and $\lambda_{data,1} = 10$ empirically.

4.4. Iteration Performance

We use alternating projection to get refined shading and reflectance. Just like what we stated in Step 5 and Step

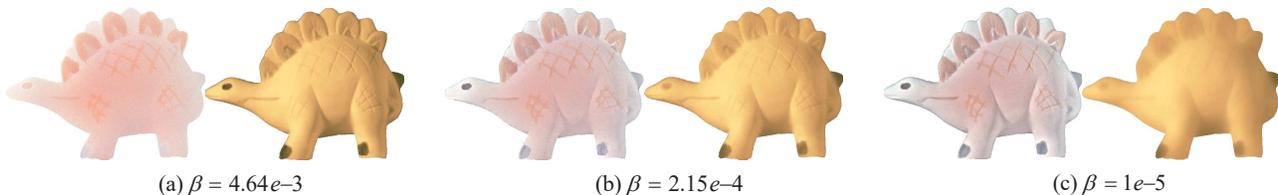


Figure 2. Estimated Reflectance and shading images using different settings for β . α is set to be 5000. (b) achieves good result while shading and reflectance overlap clearly in (a) and (c).

6 in Algorithm 1, we update shading first and then the reflectance. In each round of iteration, gradient descent method is applied to solve Eq.(10), with either \hat{S} or \hat{R} as unknowns. The iteration stops whenever it reaches the maximum iteration times 100 or when the error gradient $\nabla E < 0.01$.

We use LMSE to measure shading and reflectance images. Given the true and estimated images I and \hat{I} , [15] defined LMSE as the MSE summed over all local windows w of size $k \times k$ and spaced in steps of $k/2$:

$$\text{LMSE}_k(I, \hat{I}) = \sum_{w \in W} \text{MSE}(I_w, \hat{I}_w) = \sum_{w \in W} \|I_w - \hat{I}_w\|^2 \quad (16)$$

where $\alpha = \arg \min_{\alpha} \|I_w - \hat{I}_w\|^2$.

Fig. 3 demonstrates the iteration performance of our algorithm. The energy cost function Eq.(10) decreases with increasing iterations. Before the iteration, the color patches can still be seen clearly; while the shading image tends to be more uniform after iteration.

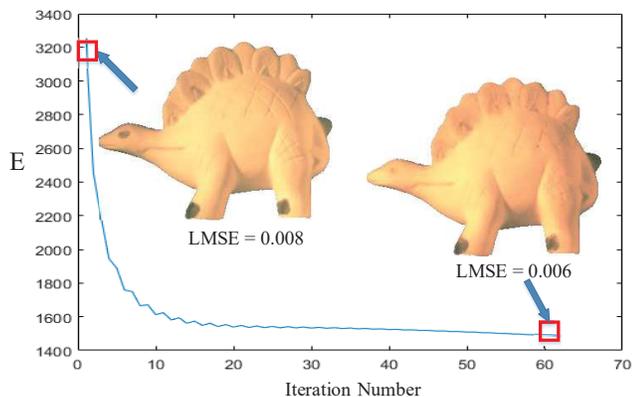


Figure 3. Iteration Performance. The algorithm converges within 100 times. Two images demonstrate the estimated shading before and after iteration.

5. Experimental Results

In this section, we provide extensive experimental validation of the proposed method. For the better visualization, we show the result in pseudo-rgb and linearly normalized the image to the range $[0, 1]$. We first show the performance of our algorithm on MIT benchmark database [15]. This is followed by our algorithm on our dataset and comparison with [12]. Finally, we test our method on Nayar [35] dataset and show visual results. Limited by the

page size, please find more dataset and results at http://cite.nju.edu.cn/MIID_dataset.html.

5.1. Experiment on MIT Benchmark Database

In order to facilitate the understanding of IID problem and illustrate the performance of our proposed method, we first tested our algorithm on MIT benchmark database [15]. Since RGB images have only 3 bands, the subspace constraint would cause significant color shift, thus we do not include the subspace constraint in this section.

Table 1. LMSE Performance on MIT benchmark database.

	CR [15]	Ours
Avg.	0.030	0.028

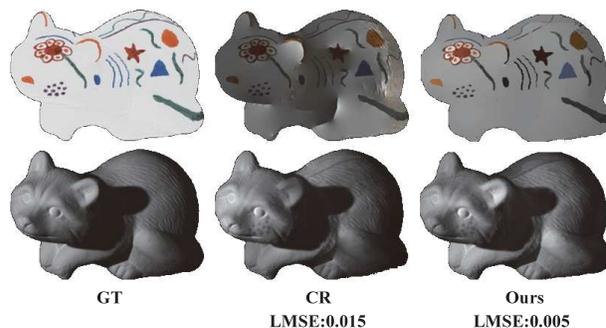


Figure 4. Visual results on MIT benchmark database. First row are reflectance images and second row are shading images.

Here GT denotes ground truth and CR indicates Color Retinex algorithm, which performed best among single image based methods in the study of [15]. From Tab. 1 and Fig. 4, it is clear that our method produces more visually pleasing decompositions with smaller LMSE.

5.2. Experiments on Proposed Dataset

We provide a benchmark dataset with ground truth for the performance evaluation of multispectral image intrinsic decomposition problem. We also compare with [12], the state-of-the-art IID algorithm in multispectral domain.

A benchmark dataset with ground-truth illumination, shading, reflectance and specularly was presented in [12] for performance evaluation of multispectral image intrinsic decomposition. Inspired by their ideas, we build up the newest multispectral intrinsic ground-truth dataset including 22 scenes under the similar environment condition, and

the light source is iodine-tungsten lamp (that interprets why shading images are a little yellowish). We apply the advanced spectrometer PMIS [9] to acquire the multispectral scenes, which could provide higher resolution in spectral data ranging from 450nm-700nm with 118 spectral channels. Compared with the dataset provided by [12], ours has 17 more scenes with higher resolution, more bands and more details, which enables the dataset to show further and potential applications in other vision researches.

Here, we evaluate our algorithm via our proposed dataset and use LMSE from the ground truth to validate our algorithm quantitatively. Compared with ground-truth, decomposition results that we achieved are desirable in terms of both the LMSE and the visual quality of the decomposed reflectance and shading results.

We display 4 examples from our dataset. The corresponding visualized RGB images for reflectance and shading are listed in Fig. 6. It is clear that our method could produce better decomposition results. In [12], superpixel-based method would lose detail information, while ours is more piecewise constant. We like to emphasize that we did all the image processing and metric computations of LMSE on down-sampled 30 out of 118 spectral channels which corresponds to the setting in [12].

Table 2. Performance statistics for dataset image

Name	SIID [12]	Ours	Name	SIID	Ours
box	0.032	0.023	ali	0.018	0.015
cup	0.016	0.012	greenpig	0.013	0.009
car	0.025	0.012	mask	0.005	0.004
bottle1	0.062	0.031	piggy	0.012	0.008
bottle2	0.005	0.008	pumpkin	0.009	0.008
bottle3	0.009	0.009	dinosaur	0.012	0.006
bus	0.030	0.031	horse	0.021	0.012
car2	0.030	0.024	kitty	0.026	0.019
dinosaur2	0.021	0.023	cap	0.027	0.024
minion	0.020	0.018	girl	0.020	0.013
plane	0.024	0.015	train	0.017	0.015
			Avg.	0.021	0.015

In table 2, we demonstrate the performance for the entire dataset. Here SIID denotes latest spectral intrinsic image decomposition in [12]. Our algorithm outperforms SIID in 19 out of 22 cases, and ours demonstrates a great improvement on average LMSE. Moreover, We note that our method is faster, more memory-friendly and is able to process larger images with more spectral bands.

In addition to evaluate the accuracy for reflectance image recovery in all spectral bands, we compare the spectral curves of selected image points from the ground truth and our algorithm. we choose patches in some scenes of our ground truth. In Fig. 5, it is obvious that our reflectance matches well with the ground truth which means we could gain accurate spectral reflectance with better performance in computation.

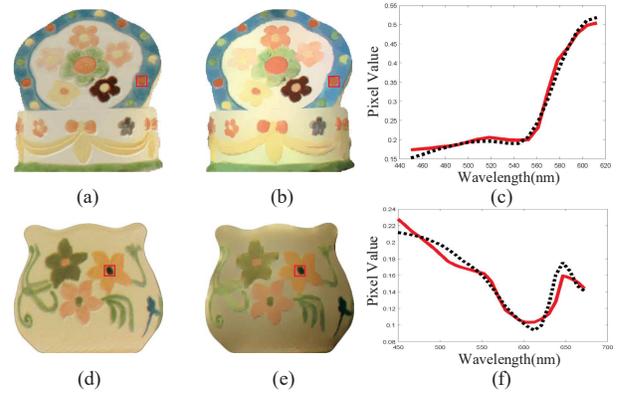


Figure 5. (a)(d) are ground truth reflectance of our dataset, (b)(e) are reflectance images of our results, and the (c)(f) are spectra curve of marked area (solid red: ours; dotted black: ground-truth).

5.3. Experiments on Nayar Multispectral Image Database[35]

To demonstrate the benefit of the subspace constraint, we compare results with and without it using Nayar Multispectral Image Database [35]. We strengthen the constraints of the 'without' one to make it solvable. In Fig. 7, the shading images with constraint are more uniform, thus reflectance images are closer to true material reflectance. From the comparison, we see that the subspace constraint helps solve the overlap between the shading and reflectance.

6. Conclusion

We have addressed the problem of the recovery of reflectance and shading from a single multispectral image captured under general spectral illumination. We have applied a subspace constraint to both the reflectance and shading space to solve the multispectral image intrinsic decomposition problem, which significantly reduce the ambiguity. Gradient descent has been used to give the initial estimation of reflectance and shading, and alternating projection method has been applied to solve the bilinear problem. Experiments on multiple datasets demonstrate that the performance of our work is better than prior works in multispectral domain.

Our work has left out constraints on global structure. Retinex theory fails to take effect when both shading and reflectance change extensively in local area. Since the high dimension of multispectral data would complicate the global constraint and bring tremendous computation cost than traditional RGB case, we hope that we will implement the global constraint with more computational efficiency (e.g. parallel design) in the near future.

Acknowledgement. This work was supported by Grant No.61627804 and 61671236 from National Nature Science Foundation of China, Grant No.BK20160634 from National Nature Science Foundation for Young Scholar of Jiangsu Province, and Grant No.0210-14380067 from Fundamental Research Funds for the Central Universities, China.

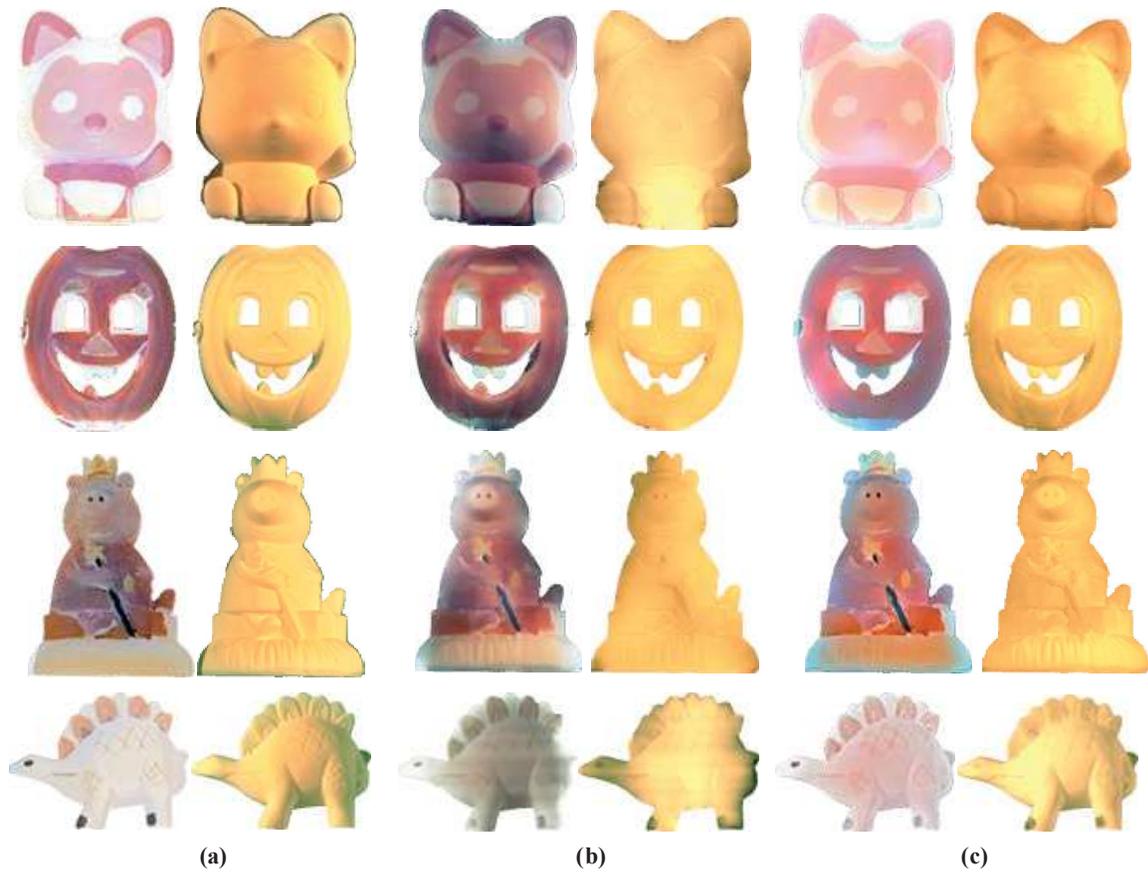


Figure 6. Results on sample images from our dataset. (a) are ground-truth reflectance and shading from our dataset, (b) are results derived from SIID [12], and (c) are our results. Note that results are rendered in RGB.

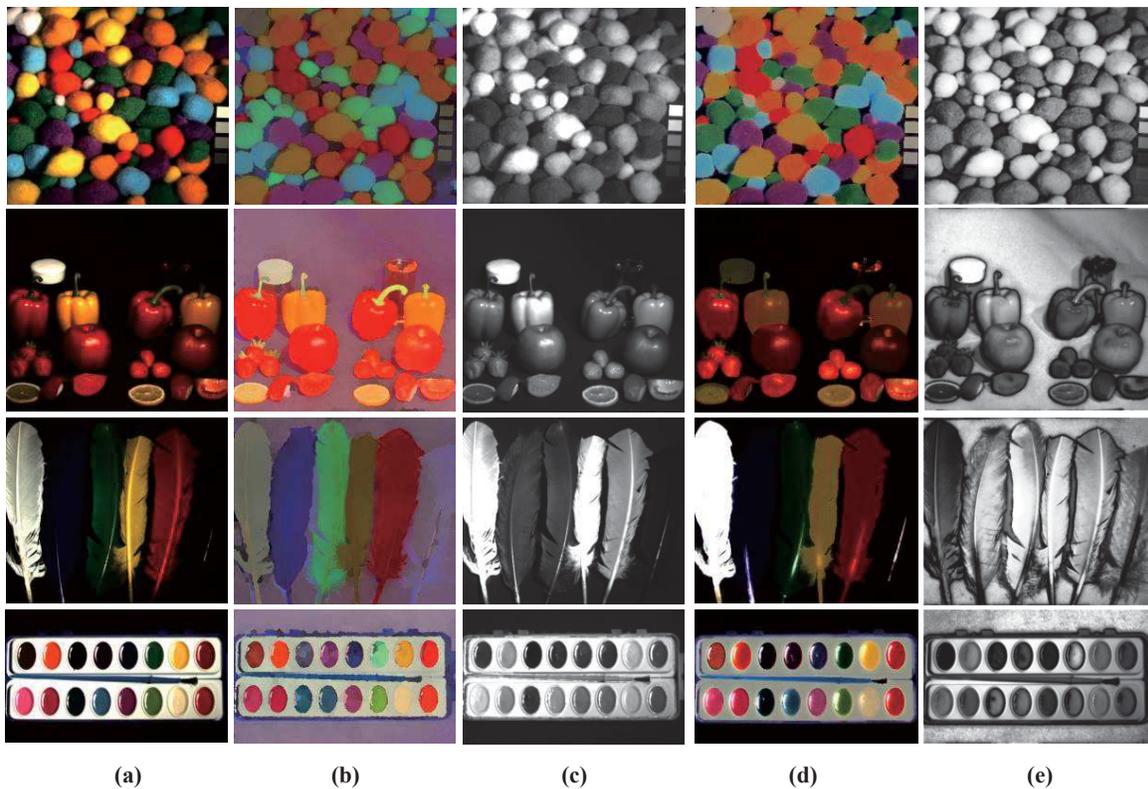


Figure 7. Comparison of decomposition without and with the subspace constraint. (a) Multispectral images (rendered in RGB) Nayar dataset [35]. (b)-(c) show the reflectance and shading components computed without the subspace constraint. (d)-(e) are with the constraint.

References

- [1] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2014. 5
- [2] J. T. Barron and J. Malik. Intrinsic scene properties from a single rgb-d image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 17–24, 2013. 2
- [3] J. T. Barron and J. Malik. Shape, illumination, and reflectance from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8):1670–1687, 2015. 2
- [4] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. *Computer vision systems*, 2, 1978. 2
- [5] M. Bell and E. Freeman. Learning local evidence for shading and reflectance. In *IEEE International Conference on Computer Vision*, volume 1, pages 670–677. IEEE, 2001. 2
- [6] S. Bell, K. Bala, and N. Snavely. Intrinsic images in the wild. *ACM Transactions on Graphics*, 33(4):159, 2014. 2, 3
- [7] S. Bi, X. Han, and Y. Yu. An l1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition. *ACM Transactions on Graphics*, 34(4):78, 2015. 2
- [8] A. Bousseau, S. Paris, and F. Durand. User-assisted intrinsic images. In *ACM Transactions on Graphics*, volume 28, page 130. ACM, 2009. 2
- [9] X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin. A prism-mask system for multispectral video acquisition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2423–2435, 2011. 5, 7
- [10] A. Chakrabarti and T. Zickler. Statistics of Real-World Hyperspectral Images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 193–200, 2011. 3
- [11] Q. Chen and V. Koltun. A simple model for intrinsic image decomposition with depth cues. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 241–248, 2013. 2
- [12] X. Chen, W. Zhu, Y. Zhao, Y. Yu, Y. Zhou, T. Yue, S. Du, and X. Cao. Intrinsic decomposition from a single spectral image. *Applied Optics*, 56(20):5676–5684, 2017. 1, 2, 3, 6, 7, 8
- [13] G. D. Finlayson, M. S. Drew, and C. Lu. Intrinsic images by entropy minimization. In *European Conference on Computer Vision*, pages 582–595. Springer, 2004. 2
- [14] B. V. Funt, M. S. Drew, and M. Brockington. Recovering shading from color images. In *European Conference on Computer Vision*, pages 124–132. Springer, 1992. 4, 5
- [15] R. Grosse, M. K. Johnson, E. H. Adelson, and W. T. Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *IEEE International Conference on Computer Vision*, pages 2335–2342. IEEE, 2009. 3, 6
- [16] J. Ho, B. V. Funt, and M. S. Drew. Separating a color signal into illumination and surface reflectance components: Theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):966–977, 1990. 2
- [17] C. P. Huynh and A. Robles-Kelly. A solution of the dichromatic model for multispectral photometric invariance. *International Journal of Computer Vision*, 90(1):1–27, 2010. 2
- [18] A. Ikari, R. Kawakami, R. T. Tan, and K. Ikeuchi. Separating illumination and surface spectral from multiple color signals. *Digitally Archiving Cultural Objects*, pages 297–321, 2008. 2
- [19] X. Jiang, A. J. Schofield, and J. L. Wyatt. Correlation-based intrinsic image extraction from a single image. In *European Conference on Computer Vision*, pages 58–71. Springer, 2010. 5
- [20] X. Kang, S. Li, L. Fang, and J. A. Benediktsson. Intrinsic image decomposition for feature extraction of hyperspectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4):2241–2253, 2015. 2
- [21] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel. A variational framework for retinex. *International Journal of Computer Vision*, 52(1):7–23, 2003. 5
- [22] P.-Y. Laffont, A. Bousseau, and G. Drettakis. Rich intrinsic image decomposition of outdoor scenes from multiple views. *IEEE Transactions on Visualization and Computer Graphics*, 19(2):210–224, 2013. 2
- [23] E. H. Land and J. J. McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971. 1, 2, 5
- [24] K. J. Lee, Q. Zhao, X. Tong, M. Gong, S. Izadi, S. U. Lee, P. Tan, and S. Lin. Estimation of intrinsic image sequences from image+ depth video. In *European Conference on Computer Vision*, pages 327–340. Springer, 2012. 2
- [25] R. Lenz, M. Österberg, J. Hiltunen, T. Jaaskelainen, and J. Parkkinen. Unsupervised filtering of color spectra. *JOSA A*, 13(7):1315–1324, 1996. data retrieved from University of Eastern Finland, Spectral Color Research Group, <http://www.uef.fi/web/spectral/-spectral-database>. 5
- [26] L. T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA A*, 3(10):1673–1683, 1986. 3, 5
- [27] A. Nadian-Ghomsheh, Y. Hassanian, and K. Navi. Intrinsic image decomposition via structure-preserving image smoothing and material recognition. *PloS one*, 11(12):e0166772, 2016. 2
- [28] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 127–136. IEEE, 2011. 2
- [29] J. P. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of munsell colors. *JOSA A*, 6(2):318–322, 1989. 3, 5
- [30] L. Shen, P. Tan, and S. Lin. Intrinsic image decomposition with non-local texture cues. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7. IEEE, 2008. 1, 2
- [31] L. Shen, C. Yeo, and B.-S. Hua. Intrinsic image decomposition using a sparse representation of reflectance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2904–2915, 2013. 1, 2

- [32] M. F. Tappen, E. H. Adelson, and W. T. Freeman. Estimating intrinsic component images using non-linear regression. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1992–1999. IEEE, 2006. [2](#), [3](#)
- [33] M. F. Tappen, W. T. Freeman, and A. E. H. Recovering intrinsic images from a single image. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 27, pages 1459–1472, 2005. [2](#), [3](#), [5](#)
- [34] Y. Weiss. Deriving intrinsic images from image sequences. In *IEEE International Conference on Computer Vision*, volume 2, pages 68–75. IEEE, 2001. [2](#)
- [35] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, 19(9):2241–2253, 2010. [3](#), [6](#), [7](#), [8](#)
- [36] Q. Zhao, P. Tan, Q. Dai, L. Shen, E. Wu, and S. Lin. A closed-form solution to retinex with nonlocal texture constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7):1437–1444, 2012. [5](#)
- [37] Y. Zheng, I. Sato, and Y. Sato. Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1779–1787, 2015. [3](#), [4](#), [5](#)
- [38] T. Zhou, P. Krahenbuhl, and A. A. Efros. Learning data-driven reflectance priors for intrinsic image decomposition. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3469–3477, 2015. [2](#)