Spectral Dichromatic Parameter Recovery from Two Views via Total Variation Hyper-priors

Filippo Bergamasco[1] Andrea Torsello[1] Antonio Robles-Kelly[2,3]

[1]Dipart. di Sci. Ambientali, Informatica e Statistica, Università Ca' Foscari Venezia, Venice Italy
[2]DATA61 - CSIRO, Tower A, 7 London Cct., Canberra, ACT 2601, Australia
[2]College of Eng. and Comp. Science, Australian National University, Canberra, Australia

Abstract. In this paper, we propose an approach for the recovery of the dichromatic model from two hyperspectral or multispectral images, *i.e.*, the joint estimation of illuminant, reflectance, and shading of each pixel, as well as the optical flow between the two views. The approach is based on the minimization of an energy functional linking the dichromatic model to the image appearances and the flow between the images to the factorized reflectance component. In order to minimize the resulting under-constrained problem, we apply vectorial total variation regularizers both to the scene reflectance, and to the flow hyper-parameters. We do this by enforcing the physical priors for the reflectance of the materials in the scene and assuming the flow varies smoothly within rigid objects in the image. We show the effectiveness of the approach compared with single view model recovery both in terms of model constancy and of closeness to the ground truth.

1 Introduction

In computer vision, the modelling and recovery of photometric parameters is a topic of pivotal importance for purposes of surface analysis and image understanding. Since the estimation of illuminant and material reflectance are mutually interdependent, the problem of recovering physically meaningful parameters that govern the image formation process is closely related to the ability to resolve the intrinsic material reflectance from their trichromatic colour images captured under varying illumination conditions. Existing methods often rely upon the use of statistics of illuminant and material reflectance or draw upon the physics-based analysis of local shading and specularity of the objects in the scene.

Statistics-based approaches often employ Bayes's rule [3] to compute the best estimate from a posterior distribution. The illuminant and surface reflectance spectra typically take the form of a finite linear model with a Gaussian basis [9], where a correlation matrix is built for a set of known plausible illuminants to characterise all the possible image colours (chromaticities) that can be observed. Contrary to these statistics-based approaches, physics-based colour constancy analyses the physical processes by which light interacts with the object surface [16,26].

Regarding specularities and shading, there have been several attempts to remove specular highlights from images of non-Lambertian objects. For instance, Brelstaff and Blake [4] used a thresholding strategy to identify specularities on moving curved objects. Narasimhan *et al.* [21] have formulated a scene radiance model for the class of "separable" Bidirectional Reflectance Distribution Functions (BRDFs). More recently, Zickler *et al.* [32] introduced a method for transforming the original RGB colour space into an illuminant-dependent colour space to obtain photometric invariants. Other alternatives elsewhere in the literature aiming at detecting and removing specularities either make use of additional hardware [22], impose constraints on the input images [18] or require colour segmentation [14].

Here, we depart from the dichromatic model so as to describe the image radiance as a combination of shading, specular highlights, surface reflectance and the illuminant power spectrum. Our multi-view dichromatic parameter recovery method separates the scene illuminant, shading and object surface reflectance by linking the reflectance of objects present in two images to the flow between between them. This is achieved by minimising the total variation of reflectance and flow, subject to the notion that the object reflectance and illuminant power spectrum across the two images should not change. This also imposes further constraints on the optical flow which are akin to those imposed upon brightness in trichromatic imagery [2]. This leads to an optimisation problem where a total variation regularization approach is used to enforce the consistency of the scene photometric parameters over an image sequence. This contrasts with previous approaches where the intersection of dichromatic planes [9,30], assumed chromaticities of common light sources [9], or structural optimisation [13] are used on single images. In [15], Kong et al. use polarised images to separate the background and reflection layers from each of the input images.

2 Contributions

The contributions of this paper are the following:

- To the best of our knowledge, this is the first approach that uses reflectance constancy across multiple images to improve the recovery of the dichromatic parameters, relating the reflectance to the optical flow between multiple images.
- We introduce a novel homographic hyper-prior for the flow similar in spirit to the affine formulation presented in [17]. This, in combination with a total variation regularization provides a natural modelling of the scene resulting in an improvement of the optical flow estimation in parallel with the improvement to the photometric parameters.
- Our method is quite general in nature and can be modified in a straightforward manner to regularise any vectorial field whose total variation under consideration is to be minimised simultaneously with other terms in an energy functional. Indeed, regularization methods have been reported in the contexts of optical flow computation [20], curvature-based surface shape representation [29] and the smoothing of stereo disparity fields [19]. We would like to stress, however, that the focus of the work presented here is the recovery of illuminant and reflectance in multispectral and hyperspectral images.
- We employ homographic hyperpriors in the total variation regularizer so as to impose a physically sound set of constraints on the solution. This not only improves the reflectance recovery results, but also delivers better localisation of the specular highlights.

– Experiments show that the approach is capable of providing a more stable recovery of illuminant, reflectance, shading and specular parameters with respect to the sate of the art. This is as our approach can achieve both, better photometric accuracy and can naturally handle the computationally challenging task of simultaneously processing a large number of wavelength indexed bands.

3 Multi-view Dichromatic Model Recovery

In this section, we present the multi-view dichromatic model energy functional which permits us, later on, to enforce consistency upon the reflectance across corresponding scene points in multiple images.

By assuming a uniform illuminant power spectrum across the scene, the dichromatic model [26] expresses the image radiance $I(\mathbf{u}, \lambda)$ at pixel location $\mathbf{u} = (u_1, u_2)$ and wavelength λ as follows:

$$I(\mathbf{u}, \lambda) = g(\mathbf{u})L(\lambda)S(\mathbf{u}, \lambda) + k(\mathbf{u})L(\lambda) \quad (1)$$

where $L(\lambda)$ and $S(\mathbf{u}, \lambda)$ are the illuminant power spectrum and surface reflectance at wavelength λ, respectively, $g(\mathbf{u})$ is the shading factor governing the proportion of diffuse light reflected from the object and $k(\mathbf{u})$ is the specular coefficient at pixel $\mathbf{u}$.

Note that the dichromatic model above assumes a single "global" illuminant power spectrum while allowing the intensity of the light to vary across the scene. This is not an overly restrictive assumption. In fact, the dichromatic model has been used extensively in colour constancy [9]. Here, we also make the assumption that $\sum_\lambda S(\mathbf{u}, \lambda)^2 = 1$. Note that this can be done without any loss of generality since the illuminant power spectrum can be normalised such that the shading factor and specular coefficients are rescaled accordingly.

3.1 Optical Flow and Reflectance Coherence

One of the main features of the dichromatic model is that the reflectance $S(\mathbf{u}, \lambda)$ is a characteristic of the object's material, being invariant to the geometry of the object and its relative position with respect to the light source and the viewer. As a consequence, it is preserved across multiple images. We model this correspondence in a two image setting by maintaining one single reflectance function on one image and relating it to the reflectance on a second image through an optical flow function $f(\mathbf{u}) = \mathbf{u}' : \Omega_1 \to \Omega_2$ which maps points from the first to the second image. This results in an energy term per each image comparing the measured irradiance $I(\mathbf{u}, \lambda)$ with the irradiance reconstructed from the model parameters using Equation (1).

Note that, for the computation of the flow, it is often assumed that the image brightness remains approximately unchanged across the two views under consideration. However, the "constant" brightness assumption applies to stereo and multiple-view settings only when the baseline is not overly wide and there is a big change in the relative angle between the objects in the scene and the illuminant direction. This assumption, however, breaks on high curvature areas introducing errors in the flow estimation as well as in the estimation of the reflectance about specular spikes and lobes.

Another problem in the recovery of the parameters is due to the fact that for highly specular pixels, the reflectance information is effectively lost at capture, *i.e.* $g(\mathbf{u})L(\lambda)$ $S(\mathbf{u},\lambda) \approx 0$. For this reason, throughout the paper, we make use of the multiplicative gating function

$$W(\mathbf{u}) = \exp\left(-\tau \| I(\mathbf{u},\lambda) - \mathcal{P}(I(\mathbf{u},\lambda)) \|\right) \tag{2}$$

where $\mathcal{P}(I(\mathbf{u},\lambda))$ is the projection of the image radiance $I(\mathbf{u},\lambda)$ onto the dichromatic plane [8] spanned by the radiance over the neighbourhood about pixel location $\mathbf{u}$. The dichromatic plane can computed using SVD [24].

The gating function above reflects the observation that, as the deviation of the image radiance from the dichromatic plane increases, the diffuse reflection decreases in importance [8]. Therefore, the function $W(\mathbf{u})$ can be viewed as a weight in the illuminant and reflectance recovery error. Further, $W(\mathbf{u})$ decreases in value for increasingly specular pixels. This is in accordance with the dichromatic plane formalism used to define $W(\mathbf{u})$, which implies that, for specular highlights, the gating function tends to zero, *i.e.* the gating function and the specular coefficient are nearly orthogonal with respect to each other. Hence, using this weighting function, the contribution of the specular pixels to the energy functional is negligible. As a result, we remove the specular coefficient $k(\mathbf{u})$ from further consideration for purposes of our optimisation approach and, instead, compute it analytically at the end of the process, once the reflectance, illuminant power spectrum, and shading are in hand.

Under these assumptions, we obtain the following energy terms comparing measured and reconstructed irradiance:

$$E_{\mathrm{DI}_1} = \int_{\Omega_1} W_1(\mathbf{u})^2 \sum_{\lambda} \big(I_1(\mathbf{u},\lambda) - L(\lambda)g_1(\mathbf{u})S(\mathbf{u},\lambda)\big)^2 d\mathbf{u}$$

$$E_{\mathrm{DI}_2} = \int_{\Omega_1} W_2\big(\mathbf{u}'\big)^2 \sum_{\lambda} \big(I_2\big(\mathbf{u}',\lambda\big) - L(\lambda)g_2\big(\mathbf{u}'\big)S(\mathbf{u},\lambda)\big)^2 d\mathbf{u}$$

where the subscript indicate the index for either of the two images. Note that, even for the term related to the second image, the integration is performed over the domain Ω_1 of the first image whereby the relations with Ω_2 is always mediated through the flow f.

3.2 Total Variation Regularization

Our goal is to minimize the energy terms over the flow f and the dichromatic model parameters. We tackle the under-determination of the problem we by adding a regularization term to the energy functional above. The *Total Variation* (TV) of a function $\phi : \mathbb{R}^m \supseteq \Omega \rightarrow \mathbb{R}^n$ is an operator defined as

$$\mathrm{TV}(\phi) = \sup_{p_1,\dots,p_m} \left\{ \int_{\Omega} \sum_{i=1}^{n} \phi_i(\mathbf{x}) \nabla \cdot p_i(\mathbf{x})\, d\mathbf{x} \;:\; p_1,\dots,p_m \in \mathcal{C}^1(\Omega,\mathbb{R}^n) \right\} \tag{3}$$

where $\mathcal{C}^1(\Omega,\mathbb{R}^n)$ is the set of continuously differentiable functions from Ω to $\mathbb{R}^n$, and $p_1,\dots,p_m$ satisfy $\sum_{i=1}^{m} \|p_i(x)\|^2 \leq 1$ everywhere except at most in a subset of measure 0. Further, if ϕ is a differentiable function, the TV assumes the equivalent form

$$\mathrm{TV}(\phi) = \int_{\Omega} ||D\phi(x)||_2 \, dx \,, \tag{4}$$

where $D\phi$ is the differential or Jacobian matrix of ϕ and $|| \cdot ||_2$ denotes the Frobenius norm.

Used as a regularizer, TV privileges piecewise constant solutions. For this property, it has found a multitude of applications ranging from image processing restoration [25], to segmentation [23], to the estimation of the optical flow [31,7]. In our proposed formulation we adopt TV to impose smoothness priors both on the reflectance and flow estimates. The reflectance component is assumed to be constant over image patches of uniform material, thus TV is naturally applicable to S, seen as a function from Ω_1 to $\mathbb{R}^\ell$ where ℓ is the number of spectral bands.

For the flow, however, there is no reason to assume a piecewise constant model. Most approaches in the literature opt to express the flow as a displacement $f(\mathbf{u}) = \mathbf{u} + T(\mathbf{u})$ where the displacement is regularized, resulting in a piecewise uniform translation. Here we opt for a higher order smoothness priors, and we compare the use of an affine prior, similar to one proposed in [17], and an homographic prior, which to the best of our knowledge has never been used before.

In the affine model we assume the displacement to be locally affine: $f(\mathbf{u}) = \mathbf{u} + A(\mathbf{u})\mathbf{u}$, where

$$A(\mathbf{u}) = \begin{pmatrix} a_1(\mathbf{u}) & a_2(\mathbf{u}) & a_3(\mathbf{u}) \\ a_4(\mathbf{u}) & a_5(\mathbf{u}) & a_6(\mathbf{u}) \\ 0 & 0 & 0 \end{pmatrix} \tag{5}$$

while for the homographic model we assume the full coordinate transformation to be projective: $\lambda f(\mathbf{u}) = H(\mathbf{u})\mathbf{u}$ where λ is a scaling factor and $H = (h_{ij})$ is in the special linear group $\mathrm{SGL}(3)$, *i.e.*, the group of 3×3 real matrices with unit determinant. Both models can be seen as capturing view transformation of locally planar patches. The homographic model does so exactly assuming the image to follow the pinhole model, while the affine model approximates it assuming a so called "weak camera model". Under these assumptions, the hyperparameters defining the entries of A and H can be assumed to be piecewise constant within such patches. In the rest of the paper we will use Θ and $\mathrm{Dom}(\Theta)$ to refer to the flow hyperparameters and their domain when these can be indifferently the affine of the homographic hyperpriors, and A or H when we want to specify models used. Note that $\mathrm{Dom}(\Theta)$ is $\mathbb{R}^6$ for the affine model and $\mathrm{SGL}(3)$ for the homographic one.

Finally, we perform a convex relaxation of the total variation functional [5] transforming the TV regularized optimization problem $\min_\phi E(\phi) + TV(\phi)$ into the relaxed problem

$$\min_{\phi,\phi_{\mathrm{TV}}} E(\phi) + \int \frac{||\phi - \phi_{\mathrm{TV}}||^2}{\delta} + TV(\phi_{\mathrm{TV}}) \,. \tag{6}$$

While the size increases with the addition of the auxiliary function ϕ_{TV}, assuming $E(\phi)$ convex, the formulation becomes convex for $\delta > 0$ and converges to the original variational problem as $\delta \to 0$.

3.3 Multi-view Dichromatic Functional

Assembling the data fidelity terms and the regularizers, we obtain the energy Multi-view dichromatic functional

$$
\begin{aligned}
E = \alpha\left(\mu E_{\mathrm{DI}_1} + (1-\mu)E_{\mathrm{DI}_2}\right) + \rho_S \int_{\Omega_1} \frac{||S(\mathbf{u}) - S_{\mathrm{TV}}(\mathbf{u})||^2}{\delta_S}\, d\mathbf{u} \\
+ \rho_S \int_{\Omega_1} ||DS_{\mathrm{TV}}(\mathbf{u})||_2 \, d\mathbf{u} + \rho_f \int_{\Omega_1} \frac{||\Theta(\mathbf{u}) - \Theta_{\mathrm{TV}}(\mathbf{u})||_2^2}{\delta_f}\, d\mathbf{u} \\
+ \rho_f \int_{\Omega_1} ||D\Theta_{\mathrm{TV}}(\mathbf{u})||_2 \, d\mathbf{u}
\end{aligned}
\tag{7}
$$

which is then minimized over S, Θ, L, g_1, g_2, S_{TV}, and Θ_{TV} subject to $\Theta \in \mathrm{Dom}(\Theta)$, to obtain simultaneous flow estimation and joint factorization of the dichromatic model over the two spectral images. Here α, ρ_S, and ρ_f are constants balancing the data fidelity and regularization terms, while $\mu \in [0;1]$ is used to limit the effect that errors in the estimation of the flow can have in the dichromatic factorization originating form the second image. Note that, as mentioned earlier, due to the $W(\mathbf{u})k(\mathbf{u})$ orthogonality we can eliminate the minimization over k, and recover the specular coefficient after the optimization from the optimal illuminant, reflectance, and shading with the relation $k(\mathbf{u}) = \frac{1}{\ell}\sum_\lambda \frac{I(\mathbf{u},\lambda)}{L(\lambda)} - g(\mathbf{u})S(\mathbf{u},\lambda)$.

4 Implementation and Discussion

4.1 Minimization Process

To optimize E we adopt an alternating minimization procedure, rotating trough the following steps:

1. Minimize with respect to $L(\lambda)$, $g_1(\mathbf{u})$, and $g_2\big(f(\mathbf{u})\big)$, keeping $S(\mathbf{u},\lambda)$, Θ, $S_{\mathrm{TV}}(\mathbf{u},\lambda)$ and $\Theta_{\mathrm{TV}}(\mathbf{u})$ fixed;
2. Update $S(\mathbf{u},\lambda)$ and Θ through a projected gradient descent step, keeping all other variables fixed;
3. Minimize the total variation terms to obtain a new estimate of $\Theta_{\mathrm{TV}}(\mathbf{u})$ and $S_{\mathrm{TV}}(\mathbf{u})$.

For the first step, we differentiate E with respect to $g_1(\mathbf{u})$ and $g_2\big(f(\mathbf{u})\big)$ and set both equations to zero so as to obtain

$$
g_1(\mathbf{u}) = \frac{\sum_\lambda I_1(\mathbf{u},\lambda)S(\mathbf{u},\lambda)L(\lambda)}{\sum_\lambda S(\mathbf{u},\lambda)^2 L(\lambda)^2} \qquad g_2\big(f(\mathbf{u})\big) = \frac{\sum_\lambda I_2\big(f(\mathbf{u}),\lambda\big)S(\mathbf{u},\lambda)L(\lambda)}{\sum_\lambda S(\mathbf{u},\lambda)^2 L(\lambda)^2}\,. \tag{8}
$$

Similarly, we differentiate $E = \mu E_{\mathrm{DI}_1} + (1-\mu)E_{\mathrm{DI}_2} + \text{const.}$ with respect to $L(\lambda)$ and set the derivative to zero, which yields

$$
L(\lambda) = C_1 \frac{\int_{\Omega_1} S(\mathbf{u},\lambda)\Delta_I(\mathbf{u},\lambda)\, d\mathbf{u}}{\int_{\Omega_1} S(\mathbf{u},\lambda)^2 \Delta_S(\mathbf{u},\lambda)\, d\mathbf{u}}\,, \tag{9}
$$

where

$$\begin{aligned}
\Delta_I(\mathbf{u},\lambda) &= \mu W_1(\mathbf{u})^2 I_1(\mathbf{u},\lambda) g_1(\mathbf{u}) + (1-\mu) W_2\big(f(\mathbf{u})\big)^2 I_2\big(f(\mathbf{u}),\lambda\big) g_2\big(f(\mathbf{u})\big)\\
\Delta_S(\mathbf{u},\lambda) &= \mu W_1(\mathbf{u})^2 g_1(\mathbf{u})^2 + (1-\mu) W_2\big(f(\mathbf{u})\big)^2 g_2\big(f(\mathbf{u})\big)^2
\end{aligned}$$

and C_1 is a normalizing constant satisfying $\sum_\lambda L(\lambda)^2 = 1$.

Hence, for the first step of the optimization process, we find the global optimum of E with respect to $g_1(\mathbf{u})$, $g_2\big(f(\mathbf{u})\big)$, and $L(\lambda)$ by alternating Equations (8) and (9). Note that, while we are estimating $g_1(\mathbf{u})$ in the regular lattice of the first image, we also do so for $g_2\big(f(\mathbf{u})\big)$ through f. This means that, in a discrete image setting, the estimated values of the second image's shading factor are not aligned with that image's regular lattice, but are shifted according to the flow f.

For the second step, we compute the gradient of E with respect to the reflectance $S(\mathbf{u},\lambda)$ and the hyper-parameter $\Theta(\mathbf{u})$. Further, in the case of the homographic hyper-parameter H, we project the gradient onto the tangent plane of $\mathrm{SGL}(3)$ before taking the gradient step and then reproject the updated H onto $\mathrm{SGL}(3)$. The constraint defining $\mathrm{SGL}(3)$ is $C = \det(H) - 1 = 0$, from which we get the projection of the gradient on the tangent space as

$$\partial_H^{\|} E = \partial_H E - \frac{\partial_H C^T \partial_H E}{\partial_H C^T \partial_H C}\partial_H C\,. \tag{10}$$

The reprojection of the updated hyperparameter H is obtained by dividing it by the cubic root of its determinant.

Note that the data fidelity term only depends of $f(\mathbf{u})$, thus, using the chain rule for the data fidelity term only, we can write

$$\partial_{\Theta(\mathbf{u})} E = (\partial_{\mathbf{u}'} E)\,(\partial_{\Theta(\mathbf{u})}\mathbf{u}') + \rho_f \frac{\Theta(\mathbf{u}) - \Theta_{\mathrm{TV}}(\mathbf{u})}{\delta_f}\,. \tag{11}$$

The gradient with respect to the flow can be computed easily in terms of the dichromatic parameters

$$\begin{aligned}
\partial_{\mathbf{u}'} E = \alpha(1-\mu)\Bigg[& E_{\mathrm{DI}_2}(\mathbf{u}')\partial_{f(\mathbf{u})} W_2\big(\mathbf{u}'\big) + 2W_2\big(\mathbf{u}'\big)\sum_\lambda \Bigg(I_2\big(\mathbf{u}',\lambda\big)\\
& - L(\lambda) g_2\big(\mathbf{u}'\big) S(\mathbf{u},\lambda)\Bigg)\cdot\Big(\partial_{\mathbf{u}'} I_2\big(\mathbf{u}',\lambda\big) - L(\lambda) S(\mathbf{u},\lambda)\partial_{f(\mathbf{u})} g_2\big(\mathbf{u}'\big)\Big)\Bigg]\,.
\end{aligned} \tag{12}$$

For the affine model, the derivative of the flow with respect to the hyperparameters is a linear function:

$$\partial_A \mathbf{u}' = \partial_{(a_1,\ldots,a_6)} f(\mathbf{u}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \big(u\ v\ 1\big) \tag{13}$$

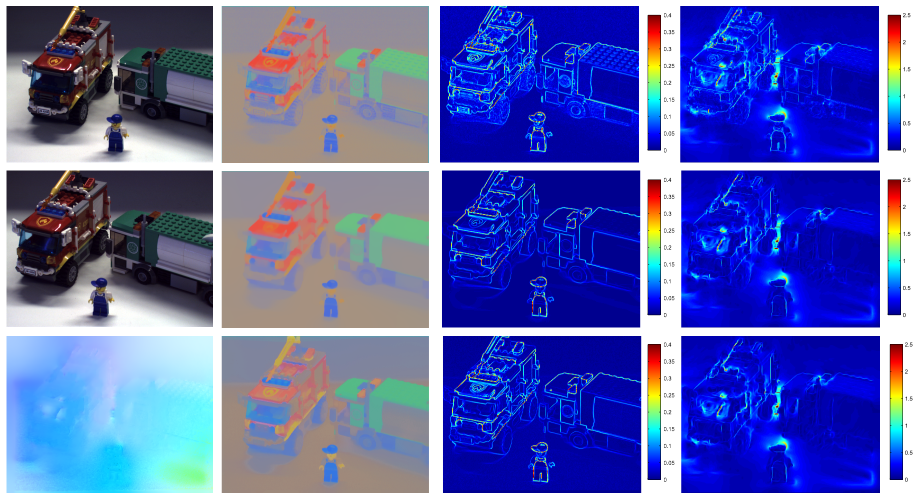

Fig. 1. Sample image pair showing the effect of the priors on the regularized parameters. First column: input image pair and optical flow. Second column: Reflectance value computed by H&RK (top) and by our approach with affine and homographic hyper-priors (center and bottom). Third column: Reflectance norm magnitude computed by H&RK (top) and by our approach with affine and homographic hyper-priors (center and bottom). Last column: Forbenious norm of the differential of the initial (top) affine and homographic hyper-parameters (center and bottom).

where $A \otimes B$ is the Kronecker product of matrices A and B and $\mathbf{u} = (u, v)^T$. For the homographic model we have:

$$\begin{aligned}\partial_H \mathbf{u}' &= (\frac{\partial \mathbf{u}'}{\partial h_{11}}, \dots, \frac{\partial \mathbf{u}'}{\partial h_{13}}, \frac{\partial \mathbf{u}'}{\partial h_{21}}, \dots, \frac{\partial \mathbf{u}'}{\partial h_{33}}) \\ &= \frac{1}{\zeta}\begin{pmatrix} 1 \; 0 \; -(h_{11}u + h_{12}v + h_{13}) \\ 0 \; 1 \; -(h_{21}u + h_{22}v + h_{23}) \end{pmatrix} \otimes \begin{pmatrix} u \; v \; 1 \end{pmatrix}\end{aligned}$$

where $\zeta = h_{31}u + h_{32}v + h_{33}$.

Furthermore, the energy gradient with respect to reflectance can be expressed as:

$$\begin{aligned}\partial_{S(\mathbf{u},\lambda)} E = &- 2\alpha\mu g_1(\mathbf{u}) W_1(\mathbf{u}) L(\lambda) \left(I_1(\mathbf{u}, \lambda) - g_1(\mathbf{u}) L(\lambda) S(\mathbf{u}, \lambda)\right) \\ &- 2\alpha(1-\mu) g_2\big(\mathbf{u}'\big) W_2\big(\mathbf{u}'\big) L(\lambda) \Big(I_2\big(\mathbf{u}', i\big) - L(\lambda) g_2\big(\mathbf{u}'\big) S(\mathbf{u}, \lambda)\Big) \\ &+ 2\rho_S \frac{S(\mathbf{u}, \lambda) - S_{\mathrm{TV}}(\mathbf{u}, \lambda)}{\delta_S}.\end{aligned}$$

We approximate $\partial_{\mathbf{u}'} W_2\big(\mathbf{u}'\big)$ and $\partial_{\mathbf{u}'} I_2\big(\mathbf{u}'\big)$ with central finite differences that are pre-computed at the beginning of the optimization process. As we mentioned before, obtaining $\partial_{f(\mathbf{u})} g_2\big(\mathbf{u}'\big)$ is not straightforward since we never optimize $g_2(\mathbf{u})$ in the regular lattice of the second image but only its representation warped to the first image through the flow $f(\mathbf{u})$. However, by the chain rule, we have:

$$\frac{\partial}{\partial \mathbf{u}} g_2\big(f(\mathbf{u})\big) = \frac{\partial g_2\big(f(\mathbf{u})\big)}{\partial f(\mathbf{u})} \cdot \frac{\partial f(\mathbf{u})}{\partial \mathbf{u}}\,, \tag{14}$$

from which we have

$$\frac{\partial g_2\big(f(\mathbf{u})\big)}{\partial f(\mathbf{u})} = \frac{\partial}{\partial \mathbf{u}} g_2\big(f(\mathbf{u})\big) \left(\frac{\partial f(\mathbf{u})}{\partial \mathbf{u}}\right)^{-1} \tag{15}$$

or, equivalently

$$\nabla_{f(\mathbf{u})} g_2\big(f(\mathbf{u})\big) = \big(D_{\mathbf{u}} f(\mathbf{u})^T\big)^{-1} \nabla_{\mathbf{u}} g_2\big(f(\mathbf{u})\big)\,, \tag{16}$$

where both terms $\nabla_{\mathbf{u}} g_2\big(f(\mathbf{u})\big)$ and $D_{\mathbf{u}} f(\mathbf{u})$ are computed with standard central differences from $g_2\big(f(\mathbf{u})\big)$ and $f(\mathbf{u})$ respectively.

Finally, for the third optimization step, we follow the fast iterative method proposed by Bresson and Chan [5].

4.2 Initialization

Note that our approach relies on an initial estimate of the flow $f(\mathbf{u})$, illuminant power spectrum and specular highlights. This is since, if the illuminant power spectrum and the specular coefficient is known, the reflectance and the shading factor can be obtained via algebraic manipulation and normalisation operations [12]. Indeed, there are a number of methods elsewhere in the literature that can be used to obtain these initial estimates. Here, we use the method in [28] to recover the image highlights and that in [9] for the recovery of the initial estimate of the illuminant power spectrum.

For the optical flow, we avoid the common coarse-to-fine-approaches, proposing to rather exploit a small set of initial sparse matches as a starting point for the flow optimisation. This is a similar approach to that used in recent works by Leordeanu *et al.*[17] or Brox and Malik [6] which are proven to deal with very large displacements. To this end, we compute a small set of reliable sparse matches from an image pair following the method in [1] and making use of *SURF* features extracted from the initial shading factor. We modified the original pay-off function to include a similarity term that weights the angular error of the reflectance spectra among two matches. As a consequence, we are better able to select a good set of inliers without epipolar filtering, which is not a feasible option if the scene objects are allowed to move.

We use these sparse matches to get an initial estimate of the flow around a limited set of points in our optimization domain, where we have designed an energy functional composed by a data term and a simple $L2$ regularizer given by

$$E_f = \alpha \int_{\Omega} D(\mathbf{u}) H(\mathbf{u}) \left[\mu_1 Es(\mathbf{u}) + \mu_2 Er(\mathbf{u})\right] + ||\partial_{\mathbf{u}} T(\mathbf{u})||_2^2 d\mathbf{u} \tag{17}$$

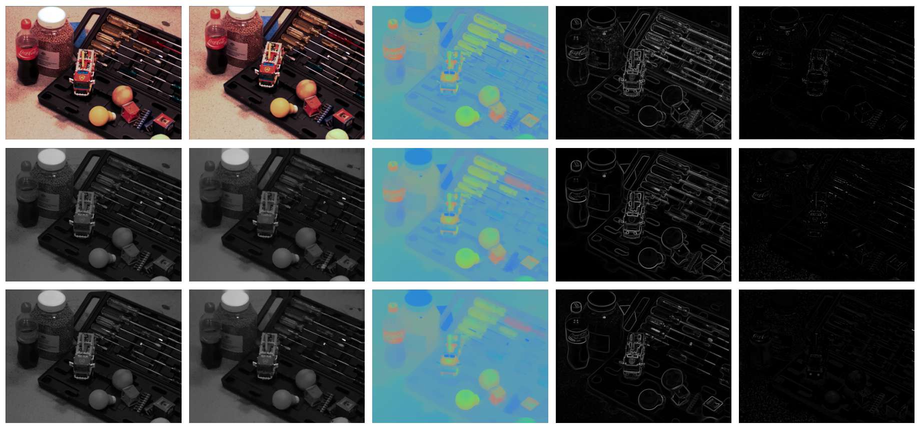

Fig. 2. Qualitative example of the results obtained for Scene 4. Top row: Initial input pairs, reflectance value, reflectance gradient magnitude and specular factors computed by H&RK. Central and Bottom rows: Shading, reflectance value, reflectance gradient magnitude and specular factors computed by the proposed method with the affine and homographic hyper-priors respectively.

with

$$Es(\mathbf{u}) = [g_1(\mathbf{u}) - g_2\big(\mathbf{u} + T(\mathbf{u})\big)]^2 \tag{18}$$

$$Er(\mathbf{u}) = e^{-\sum_\lambda S_1(\mathbf{u},\lambda) S_2\big(\mathbf{u}+T(\mathbf{u})),\lambda\big)} \tag{19}$$

$$D(\mathbf{u}) = e^{-\frac{1}{\sigma}\min_{m \in M} \|\mathbf{u}-m\|} \tag{20}$$

$$H(\mathbf{u}) = \gamma \frac{\sum_\lambda \|\partial_\mathbf{u} S_1(\mathbf{u},\lambda)\|^2}{\max_{\mathbf{u}'} \sum_\lambda \|\partial_{\mathbf{u}'} S_1(\mathbf{u}',\lambda)\|^2} + 1 \tag{21}$$

where γ, μ_1 and μ_2 are constants, σ is the radius of the spatial weighting term and we have written $S_i(\mathbf{u},\lambda)$ to imply that the reflectance corresponds to the i^{th} view under consideration, i.e. $i = \{1,2\}$.

In the expression above, the data term accounts for both the photometric Es and material Er consistency between the two images trough an $L2$-norm penalty function. The spatial weighting term D moderates the effect of the $L2$-regularizer with respect to the data term as a function of the distance from the closest match in the initial set M whereas (H) is used to allow discontinuities in the proximity of edges. Here, we tackle the minimization of the functional above as a standard variational problem by solving the set of associated Euler-Lagrange equations [10] and have set all constants to unity.

4.3 Effect of the regularization terms

Recall that the total variation hyper-prior regularization term was introduced to enforce the patch-wise uniform material assumption and the locally uniform flow assumption formalized in terms of locally-homographic transformation. Figure 1 shows the effect of the priors on the regularized parameters on a sample hyperspectral image pair. The left-most column shows the input color image pair (we show the pseudocolour obtained using the colour matching functions of Stiles and Burch [27] over the 30 bands in the visible range as delivered by the camera) and the initial flow. The second column shows the reflectance (again, in pseudocolour) as returned by [12] (H&RK) and as optimized by our process with the affine and homographic priors respectively in the first, second and third row. The third column shows the gradient magnitude of the reflectance for the same three methods, while the last column shows the Frobenius norm of the differential of the hyper-parameters A and H at initialization and after optimization.

From the figure, it is clear that the algorithm is capable of clustering together regions of uniform material that had significant variation in the estimated reflectance with H&RK. For example, look at the gradient magnitude of the reflectance in areas like the roof of the truck on the right or the wheels of the trick on the left. In both cases the materials are uniform and thus should exhibit uniform reflectance, but the wide variation in shading leaks into variations in the reflectance estimated with H&RK, on the other hand, our approach strongly reduces the variation in reflectance, while maintaining sharp variations across different materials.

Also note that the regularization of the hyper-parameters significantly improve the details captured by the flow. For instance, the flow around the logos on the two trucks correspond to a pure change in material and should not have any effect on the flow. However, the edges of the logos are clearly visible in the gradient magnitude of the flow hyper-parameters at initialization, which indicates a leakage of information from the estimated reflectance to the estimated flow. After optimization, not only is the flow generally more uniform, with high gradient mostly in correspondence with depth discontinuities or occluded pixels, but the boundaries of the logos vanish almost completely. Indeed, from the reflectance gradient in the 3^{rd} column, we can see many more shading artefacts and edge ghosting effects in the results yielded by the H&RK alternative than in those obtained with our approach. As expected, the homographic hyper-prior performs better than the affine one as it captures the assumption of locally planar patches. This is clear by observing the yellow logo on the truck. The affine case (central row) suffers a loss of contrast which is almost negligible for the homographic case while still effectively clustering uniform patches of similar material.

5 Experiments

For purposes of comparison, we have used the method in [12]. Our choice hinges in the fact that the alternative is aimed at processing imaging spectroscopy data based upon the dichromatic model. Moreover, the method in [12] is an optimisation approach. Both our method and the alternative have been initialised using the same estimates of the illuminant power spectrum and specular highlights.

For the experiments shown in this section, we have used four image sequences acquired using an uncalibrated multispectral camera delivering six channels in the visible spectrum *i.e.* six wavelength indexed bands in the range $430 - 680nm$ with $50nm$ steps, and one in the near-infrared at $950nm$. It is worth noting in passing that our method can be easily applied to data comprising any number of wavelength bands, as shown in Figure 1, where the images comprised 30 bands. Each of our image sequences

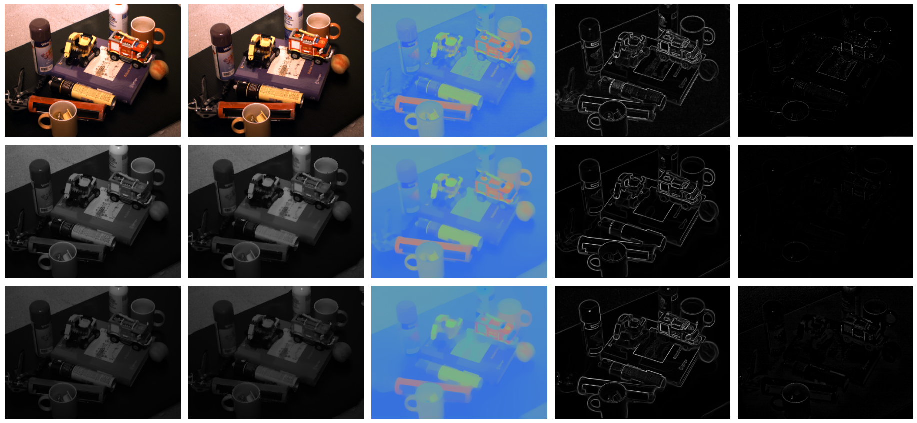

Fig. 3. Qualitative example of the results obtained for Scene 5. Top row: Initial input pair, reflectance value, reflectance gradient magnitude and specular factors computed by H&RK. Central and Bottom rows: Shading, reflectance value, reflectance gradient magnitude and specular factors computed by the proposed method with the affine and homographic hyper-priors respectively.

here comprises 10 frames, depicting scenes containing a wide variety of objects made of different materials and depicting a wide variety of shapes. Each of these scenes is illuminated by different lights, spanning artificial sunlights, tungsten and incandescent lamps. For each of these, the ground truth illuminant power spectrum has been acquired using a LabSphere Spectralon calibration target. For our dataset, we have computed reflectance images for groundtruthing purposes following the procedure in [11]. All our pseudocolour images have been obtained using the colour matching functions of Stiles and Burch [27].

In Figures 2 and 3 we present some qualitative results comparing the proposed method against H&RK for two sample scenes in our dataset. In the first row, a pseudocolour image pair obtained in the same way as in [12] is shown together with the reflectance, reflectance gradient magnitude and specular factor estimated by H&RK. In the second and the third row the output of our method for affine and homographic hyper priors respectively is shown together with the estimated shading factor. Note how the total variation regularisation process in our approach has improved the reflectance estimate by removing artefacts arising from the surface geometry. The specular highlights delivered by our method are in better accordance with the input imagery as can be appreciated by observing the screwdriver box, the jars and the cups present in the scenes. Overall, both the regularizer performs better than H&RK producing similar results. However, a better estimation of the flow given by the homographic hyperprior improves the reflectance contrast and specular highlights localization.

In Figures 4 and 5 we illustrate the results yielded by our method and the alternative regarding the recovery of the reflectance and the illuminant power spectrum. To this end, the left-hand side of Figure 4, we plot the illuminant delivered by H&RK and our method superimposed over the ground-truth (red-line). In the panel, the top trace depicts the spectrum whereas the bottom bar plot corresponds to the standard deviation of the

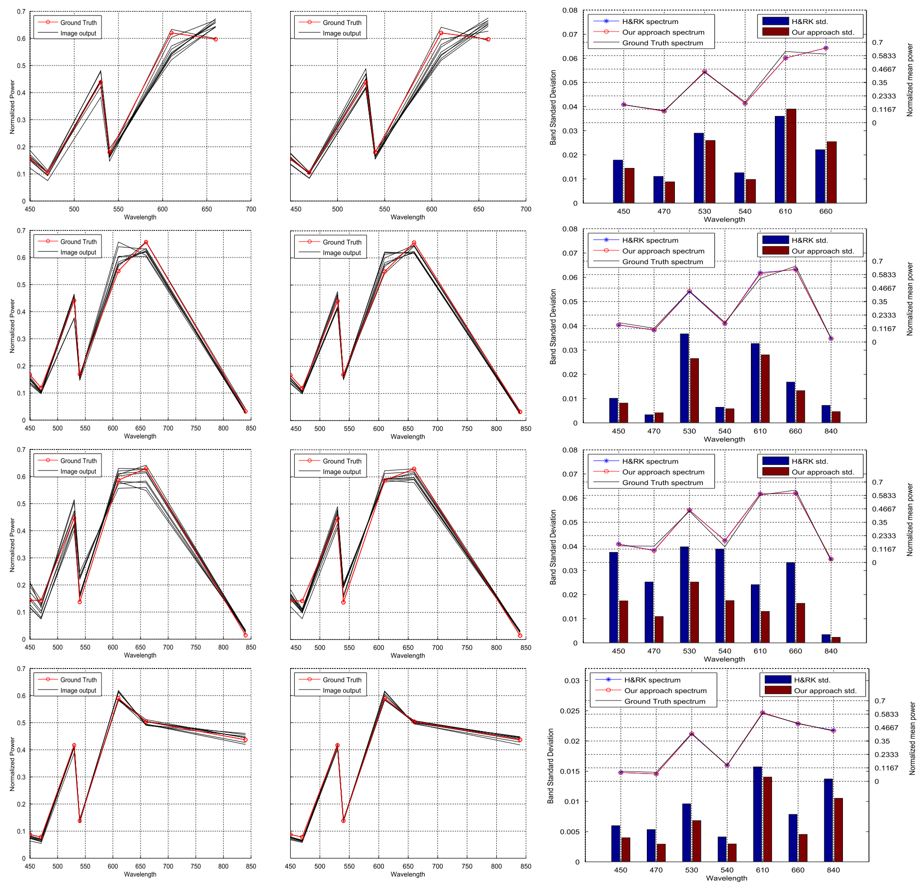

Fig. 4. Illuminant power spectra for each scene. First and second column: power spectrum for each scene image as computed by H&RK and our approach respectively. Third column: average spectrum with the standard deviation for each band.

illuminant per band over the corresponding image sequence. Note that the standard deviation for the illuminant power spectrum is much lower for our method. This is also the case for the reflectance. In Figure 5 we show the reflectance for four colour tiles on an XRite colour checker placed in one of our scenes. For the sake of clarity of presentation, the figure shows a close up of the color checker and, in a fashion similar to Figure 4, the spectrum as a trace at the top of the plots with the standard deviation at the bottom on a bar plot.

In Table 5 we show the RMS and Euclidean angular error for the illuminant power spectrum recovered by our approach and the H&RK method across all the images for the four scenes in our dataset. Note that, for both measures, our method exhibits a lower error. This is consistent with our qualitative results showed earlier. Finally, in Table 5,

	mean	st. dev.
H&RK θ	0.075242	0.01740
Our θ	0.073063	0.02032
H&RK RMS	0.028431	0.00657
Our RMS	0.027607	0.00767

Table 1. Average and standard deviation of the RMS and Euclidean angular error (θ) of the estimated reflectance inside the coloured tiles shown in Figure 5.

Scene	H&RK θ	Our θ	H&RK RMS	Our RMS
1	0.080354	0.080045	0.026777	0.026675
2	0.066695	0.055120	0.023576	0.019485
3	0.076167	0.074565	0.025383	0.024849
4	0.021691	0.020638	0.007669	0.007296

Table 2. RMS and Euclidean angular error (θ) for the illuminant recovered by our approach and the H&RK method for all the four scenes of our dataset.

we show the RMS and Euclidean angular error of the recovered reflectance averaged among each coloured tile shown in Figure 5.

6 Conclusions

In this paper, we proposed a novel method for dichromatic model recovery from a spectral image pair by means of an energy minimization that simultaneously take into account the model parameters and the flow between the images. We introduced a novel affine hyper-prior for the flow that, in combination with a Total Variation regularization, provides a natural piecewise-planar assumption of the scene under the pinhole camera

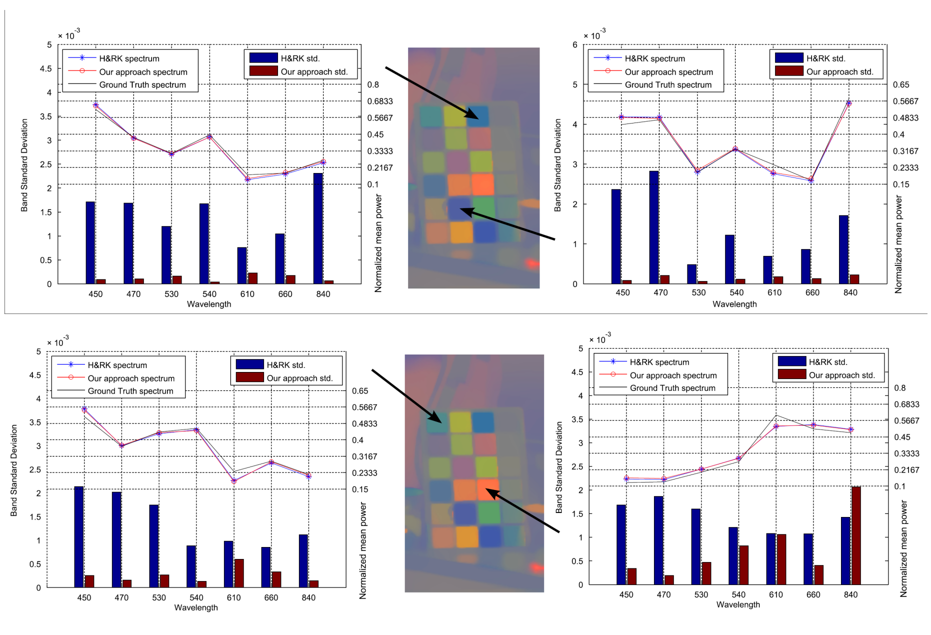

Fig. 5. Reflectance spectra and the standard deviation for each band for the pixels inside the respective color tiles.

model. The same kind of regularizer is used for the reflectance imposing the assumption that objects are composed by local patches of uniform materials. As a result, we are able to obtain a better reflectance estimation with respect to the current single-image state of the art approaches. Moreover, our approach has shown a significant lower variance while computing the illuminant spectrum over a sequence of images of the same scene. This behaviour is crucial for many applications for which a coherence of the dichromatic parameters is advisable when analysing multiple instances of the same objects involved in a sequence for which the illuminant is constant across the scene. Furthermore, qualitative results shows that the method discriminates better between the shading (i.e. the geometrical features of a surface) and the texture an object.

References

1. Albarelli, A., Rodolà, E., Torsello, A.: Imposing semi-local geometric constraints for accurate correspondences selection in structure from motion: A game-theoretic perspective. IJCV 97(1), 36–53 (2012) 9
2. Barron, J.L., Fleet, D.J., Beauchemin, S.S.: Performance of optical flow techniques. Int. Journal of Computer Vision 12(1), 43–77 (1994) 2
3. Brainard, D.H., Delahunt, P.B., Freeman, W.T., Kraft, J.M., Xiao, B.: Bayesian model of human color constancy. Journal of Vision 6(11), 1267–1281 (2006) 1
4. Brelstaff, G., Blake, A.: Detecting specular reflection using lambertian constraints. In: Int. Conference on Computer Vision. pp. 297–302 (1988) 1
5. Bresson, X., Chan, T.F.: Fast dual minimization of the vectorial total variation norm and applications to color image processing (2008) 5, 9
6. Brox, T., Malik, J.: Large displacement optical flow: Descriptor matching in variational motion estimation. IEEE TPAMI 33(3), 500–513 (March 2011) 9
7. Drulea, M., Nedevschi, S.: Total variation regularization of local-global optical flow. In: Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on. pp. 318–323 (Oct 2011) 5
8. Finlayson, G.D., Schaefer, G.: Convex and non-convex illuminant constraints for dichromatic colour constancy. In: IEEE CVPR. pp. I:598–604 (2001) 4
9. Finlayson, G.D., Schaefer, G.: Solving for colour constancy using a constrained dichromatic reflection model. IJCV 42(3), 127–144 (2001) 1, 2, 3, 9
10. Forsyth, A.: Calculus of variations. Dover books on advanced mathematics, Dover Publications (1960) 10
11. Foster, D.H., Amano, K., Nascimento, S.M.C., Foster, M.J.: Frequency of metamerism in natural scenes. J. Opt. Soc. America A 23(10), 2359–2372 (2006) 12
12. Huynh, C.P., Robles-Kelly, A.: A solution of the dichromatic model for multispectral photometric invariance. IJCV 90(1), 1–27 (2010) 9, 11, 12
13. Huynh, C.P., Robles-Kelly, A., Hancock, E.R.: Shape and refractive index recovery from single-view polarisation images. In: IEEE CVPR (2010) 2
14. Klinker, G., Shafer, S., Kanade, T.: A physical approach to color image understanding. Intl. Journal of Computer Vision 4(1), 7–38 (1990) 2
15. Kong, N., Tai, Y., Shin, J.S.: A physically-based approach to reflection separation: From physical modeling to constrained optimization. IEEE Trans. on Pattern Anal. and Mach. Intell. 36(2), 209–221 (2014) 2
16. Land, E.H., Mccann, J.J.: Lightness and retinex theory. J. Opt. Soc. Am 61, 1–11 (1971) 1
17. Leordeanu, M., Zanfir, A., Sminchisescu, C.: Locally affine sparse-to-dense matching for motion and occlusion estimation. In: IEEE ICCV (December 2013) 2, 5, 9

18. Lin, S., Shum, H.: Separation of diffuse and specular reflection in color images. In: Int. Conf. on Comp. Vision and Patt. Recognition (2001) 2
19. Marr, D., Poggio, T.: A computational theory of human stereo vision. In: Proceedings of the Royal Society of London. Series B, Biological Sciences, vol. 204, pp. 301–328 (1979) 2
20. Nagel, H., Enkelmann, W.: An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. IEEE Trans. on Pattern Analysis and Machine Intelligence 8, 565–593 (1986) 2
21. Narasimhan, S.G., Nayar, S.K.: Contrast restoration of weather degraded images. IEEE TPAMI 25, 713–724 (2003) 1
22. Nayar, S., Bolle, R.: Reflectance based object recognition. International Journal of Computer Vision 17(3), 219–240 (1996) 2
23. Pock, T., Cremers, D., Bischof, H., Chambolle, A.: An algorithm for minimizing the mumford-shah functional. In: ICCV. pp. 1133–1140. IEEE (2009) 5
24. Robles-Kelly, A., Huynh, C.P.: Imaging Spectroscopy for Scene Analysis. Springer (2013) 4
25. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D 60(1-4), 259–268 (1992) 5
26. Shafer, S.A.: Using color to separate reflection components. Color Research and Applications 10(4), 210–218 (1985) 1, 3
27. Stiles, W.S., Burch, J.M.: Interim report to the Commission Internationale de l'Éclairage Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching. Optica Acta 2, 168–181 (1955) 11, 12
28. Tan, R.T., Nishino, K., Ikeuchi, K.: Separating reflection components based on chromaticity and noise analysis. IEEE TPAMI 26(10), 1373–1379 (2004) 9
29. Terzopoulos, D.: Multilevel computational processes for visual surface reconstruction. Computer Vision, Graphics and Image Understanding 24, 52–96 (1983) 2
30. Tominanga, S., Wandell, B.A.: Standard surface-reflectance model and illuminant estimation. Journal of the Optical Society of America A (6), 576–584 (1989) 2
31. Werlberger, M., Pock, T., Bischof, H.: Motion estimation with non-local total variation regularization. In: CVPR. pp. 2464–2471. IEEE (2010) 5
32. Zickler, T., Mallick, S.P., Kriegman, D.J., Belhumeur, P.N.: Color subspaces as photometric invariants. IJCVs 79(1), 13–30 (2008) 2