Segmentation and Estimation of Spatially Varying Illumination

Lin Gu, *Student Member, IEEE,* Cong Phuoc Huynh, *Member, IEEE,* and Antonio Robles-Kelly, *Senior Member, IEEE*

Abstract

In this paper, we present an unsupervised method for segmenting the illuminant regions and estimating the illumination power spectrum from a single image of a scene lit by multiple light sources. Here, illuminant region segmentation is cast as a probabilistic clustering problem in the image spectral radiance space. We formulate the problem in an optimisation setting which aims to maximise the likelihood of the image radiance with respect to a mixture model while enforcing a spatial smoothness constraint on the illuminant spectrum. We initialise the illuminant for each pixel via a projection of the image radiance spectra onto a low-dimensional subspace spanned by a randomly chosen subset of spectra. Subsequently, we optimise the objective function in a coordinate-ascent manner by updating the weights of the mixture components and the illuminant posterior probability. We then estimate the illuminant power spectrum per pixel in a straightforward manner by applying existing colour constancy methods to each of these pixel clusters. We compare our method with a number of alternatives for the tasks of illumination region segmentation, illumination colour estimation and colour correction. Our experiments show the effectiveness of our method as applied to one hyperspectral and four trichromatic image datasets.

Index Terms

Illumination estimation, multiple illuminant segmentation, spatially varying colour constancy

I. Introduction

Detecting and estimating illuminant colours are important tasks with applications to recognition and classification based on photometric invariants [1], white balancing [2], colour correction, digital media production and graphics [3]. Despite its importance, the recovery and identification of illuminant colours

L. Gu is with the Research School of Engineering, ANU, Canberra, ACT 0200, Australia. E-mail: lin.gu@anu.edu.au.

Cong Phuoc Huynh and Antonio Robles-Kelly are with NICTA (National ICT Australia), Locked Bag 8001, Canberra, ACT 2601, Australia.

in a scene has proven a difficult task in uncontrolled real world imagery. This is mainly due to the fact that the recovery of the region-wise illuminant from a single image is an under-constrained problem [4]. As a result, existing methods often assume a uniform illumination power spectrum throughout the scene [5], [6]. Other works have made statistical assumptions on the image colours to remove dependency on material colours, such as the "grey world" assumption [7], which implies that the statistical mean colour of a scene is achromatic. The "white patch" hypothesis [8] relies on the existence of a white pixel in the scene for illuminant colour inference. Both of the methods above are special instances of the "shades of grey" method [9] and are akin to edge-based colour constancy methods [10], which assume that the statistical mean of spatial derivative across the image is only dependent on the colour of the illuminant, rather than the scene materials. Further, for colour balancing and correction, Hsu *et al.* [2] recover a set of dominant material colours to estimate the local mixture coefficients of the lights. Recently, several authors proposed methods for white balance under mixed lighting conditions using user-assisted inputs and sparse interpolation [11], [12].

These assumptions are often too restrictive for real world scenes since shadows, skylight and direct sunlight can cause significant illuminant variations across the image [13], [14]. To handle varying illumination, Finlayson *et al.* [14] mapped image chromaticities to those of standard colour patches rendered under canonical illuminants. The Retinex theory [**?**], [**?**] has also been used for the removal of local illumination taking into account the spatial changes in the image intensity. Ebner [15] employs the local space average to perform colour constancy irrespective of the illuminants used in the scene. Wang and Samaras [16] detect and estimate multiple illuminants from a single image of an object with known geometry and Lambertian reflectance using a region-based recursive least squares method. More recently, Finlayson *et al.* [17] detected the illumination colour in images using a chromagenic camera.

Unfortunately, the algorithms above do not contain a robust mechanism to distinguish illumination changes from reflectance variations. Further, the uneven nature of the shading over the surface radiance often adversely affects the stability of the results. To overcome these shortcomings, Barnard *et al.* [18] addressed the colour constancy problem by combining constraints based on the variations over the surface reflectance and illumination. To do this, they employ prior knowledge about plausible illuminants so as to characterise the image chromaticity changes caused by illumination variation. In a related approach, Gijsenij *et al.* [19] recover the illuminant in non-uniformly lit scenes by locally applying colour constancy

over a set of image patches. Bleier *et al.* [4] apply colour constancy methods for uniform illumination to each super-pixel resulting from a segmentation of the input image. Riess *et al.* [20] recover the illuminant colour per super-pixel from a single image by grouping local estimates into regions with similar illuminant colours.

In this paper, we present an unsupervised method for the segmentation of illumination regions and the estimation of the illuminant power spectrum in a scene lit by multiple light sources. The method presented here is applicable to a wide variety of scenes where the illumination power is piece-wise constant or slowly varying across image regions. To this end, we assume the illuminant colour varies smoothly in the spatial domain. This local smoothness assumption permits the enforcement of illumination consistency across the spatial domain.

Hence, we view the problem of segmenting illumination regions as a clustering one by viewing the illumination segmentation as a clustering problem with respect to a mixture model. The clustering process occurs in an iterative manner where we initialise the algorithm with an effective separation of the illuminant regions based on the projection onto random subspaces presented in [21]. Thus, we employ a kernel density estimator [22] to approximate the likelihood of the image irradiance occuring under a particular light source. In addition, the local smoothness constraint on the illuminant allows the correction of spatial discontinuities across the resulting illumination regions. Subsequently, the average illuminant for each region is computed by applying colour constancy methods used elsewhere for single light source scenes, such as Grey-World [7], Grey-Edge [10] and White-Patch [8]. Finally, a per-pixel estimation of the illuminant spectrum can be obtained as the weighted average of the illuminant per region, where the weights are the posterior probabilities of the illumiant at each pixel.

Our method confers several advantages over previous work. Firstly, the number of illuminants and their colours do not need to be pre-determined. This contrasts with other approaches elsewhere such as the work reported in [2]. In fact, one can initialise our algorithm with a sufficiently large number of illuminants and the method will eventually converge to a minimal number of distinctive light sources. Secondly, our method does not require user-intervention. Being completely unsupervised sets it apart from methods such as those in [11], [12]. In addition, unlike the work in [23], the illumination boundaries can be designated without sampling image patches. As a result, the illuminant colour is independent of the sampling strategy and the patch size. Furthermore, by estimating the illumination boundaries, we avoid the

risk of over-smoothing the illuminant spectra, which is an issue encountered by Ebner's local-averaging method [15].

The paper is organised as follows. In Section II, we provide the background as related to the radiance space used throughout the paper. Subsequently, we formulate the illuminant segmentation and estimation problem as an statistical inference one and derive a coordinate-descent optimisation strategy in Section III. We show results for our method and compare to a number of alternatives in Section IV. Finally, we conclude upon the work presented here in Section V.

II. Spectral Radiance Space

In this section, we examine the structure of the radiance spectra in a scene illuminated by multiple light sources. The idea is to do this so as to view the illuminant segmentation problem as a clustering one in the spectral radiance space. To do this, we depart from the image formation process. Consider the location u in the scene is illuminated by a power spectrum $L(\lambda)$, where λ is the sample wavelength variable. The spectral radiance $I(u,\lambda)$ reflected from that location can be expressed as

$$I(u,\lambda) = L(\lambda)S(u,\lambda), \tag{1}$$

where $S(u,\lambda)$ is the reflectance function at location u and wavelength λ.

Note that the model in Equation 1 has been used widely in the literature for illumination recovery [24]. Further, we refer to existing research on linear models of reflectance spectra of real-world materials. In an early experiment, Cohen [25] found that the spectral reflectance of 433 chips in the Munsell Book of Color lie in a subspace spanning three basis components. Subsequently, Maloney [26] analysed the fitting of Munsell color samples and naturally occuring reflectance spectra to a linear combination of two to six basis elements. Building on these experimental results, Marimont and Wandell developed a low-dimensional linear model of real-world spectral reflectance [27] where each surface reflectance $S(u,\lambda)$ is expressed as a weighted sum of wavelength-dependent basis functions $S_i(\lambda)$. This is given by

$$S(u,\lambda) = \sum_i w_i(u)S_i(\lambda), \tag{2}$$

where the basis functions can be obtained via Principal Component Analysis (PCA) [28] or Independent Component Analysis (ICA) [29].

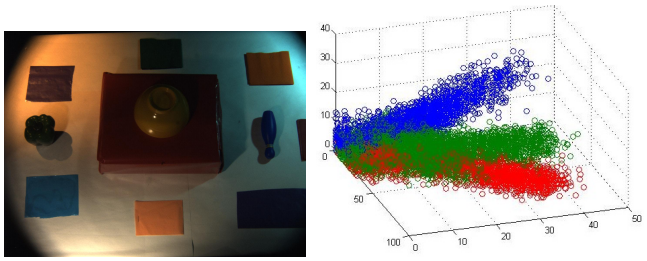


Fig. 1. The distribution of colours of a scene illuminated by three different light sources. The left panel shows the RGB image of a scene. The right panel shows three clusters of RGB colour vectors, each for an illumination region. These have been plotted in a three dimensional coordinate system where the axes correspond to the R, G and B channels.

Combining Equation 1 and 2, the scene radiance can be written as a linear combination of the basis radiance spectra as follows

$$I(u,\lambda) = \sum_i w_i(u) B_i(\lambda), \tag{3}$$

where $B_i(\lambda) = L(\lambda)S_i(\lambda)$ is the spectral radiance of the i^{th} basis component.

Since the weights $w_i(u)$ in Equation 2 are non-negative, Equation 3 implies that the radiance of all real-world materials under a fixed illuminant spectrum form a convex cone with $B_i(\cdot)$ as its vertices. This observation serves as the foundation for gamut mapping methods under uniform illumination [30]. Furthermore, here we note that each basis vector $B_i(\cdot)$ is dependent on the illuminant spectrum $L(\cdot)$. As a consequence, the coordinates of $B_i(\cdot)$ in the colour space are skewed by the direction of the illuminant spectrum and so is the orientation of the convex cone of image colours.

This is illustrated in Fig 1, where three regions in the input scene (left-hand panel) are lit by different illuminants whose hues are blue, orange and yellow, respectively. The right-hand panel shows the distribution of colour vectors of all the image pixels in the RGB colour space, which have been colour-coded according to the illuminant at each pixel. Note that these colour vectors indeed form clusters with orientations skewed by the respective illuminant. With this in mind, the multiple convex cones in the spectral radiance space can be viewed as clusters induced by distinct light sources in the radiance space. This is the basis observation for our clustering approach, allowing for the separation of pixels illuminated by each light source.

III. Illuminant Segmentation and Estimation

We now consider the problem of segmenting and estimating the spatially varying illuminant spectra in the scene. We would like to stress that, although the following derivation takes the number of scene illuminants as input, it is, in general, not a requirement for our method. In fact, the number of illuminants can be estimated making use of a pre-processing step such as that in [31]. In our implementation, the algorithm commences with a sufficiently large number of illuminants. Illuminants found to be overly similar in terms of their angular difference are then merged and their contributions to each pixel are consolidated at the end of the algorithm.

As said before, we assume that the illuminant colour vary smoothly in the spatial domain. This assumption permits the computation of each component illuminant by standard colour constancy algorithms. Such assumption also gives rise to a sampling approach which yields pixel subsets illuminated by each illuminant. Our illuminant segmentation scheme maked use of an expansion step for these subsets through a substitution strategy over several iterations.

A. Objective Function

We formulate the target function of the clustering problem as follows. Firstly, we denote the spectral radiance vector at pixel u as $\mathbf{I}(u) \triangleq (I(u,\lambda_1), I(u,\lambda_2), \ldots, I(u,\lambda_K))$, where the λ_k's, $k = 1,\ldots,K$, are the sampled wavelengths of the image. Consider a scene illuminated by M illuminants $L_1, L_2, \ldots L_M$. We denote α_m as the mixing coefficient of the cluster of pixel radiance spectra under the m^{th} illuminant. The mixing coefficients $\alpha_m, m = 1,\ldots,M$ indicate the prior probability of the m^{th} illuminant and satisfy the conditions $\sum_{m=1}^{M} \alpha_m = 1$ and $\alpha_m \geq 0, \forall m = 1,\ldots,M$. In addition, we denote the cluster of pixels illuminated by the m^{th} illuminant as Ω_m, where $\Omega_i \cap \Omega_j = \emptyset \ \forall i \neq j$, and the union of these as $\Omega \triangleq \cup_{m=1}^{M} \Omega_m$.

We cast our problem as one that finds the clusters $\Omega_1, \Omega_2, \ldots, \Omega_M$ that maximise the likelihood of the input image radiance. Here, we formulate the likelihood of the spectral radiance vector $\mathbf{I}(u)$ given the pixel clusters as

$$p(\mathbf{I}(u)|\Omega) = \sum_{m=1}^{M} \alpha_m p_m(u), \tag{4}$$

where $p_m(u) \triangleq p(\mathbf{I}(u)|\Omega_m)$ is the probability density function (PDF) of the radiance vector $\mathbf{I}(u)$ given the set of pixels Ω_m illuminated by the m^{th} illuminant.

Here, we approximate this PDF using kernel density estimation [22]. Thus, we rely on a kernel function $K_h(\cdot)$ with bandwidth h. This function aims to express the probability density of a radiance spectrum $\mathbf{I}(u)$ in terms of the distances to the radiance spectra of the sample pixels in Ω_m. Formally, the kernel density estimate of PDF $p_m(u)$ is given by

$$p_m(u) \approx \frac{1}{n_m h} \sum_{v \in \Omega_m} K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|), \tag{5}$$

where $n_m = |\Omega_m|$ is the number of sample pixels illuminated by the m^{th} illuminant and $\|\cdot\|$ is the L^2-norm of the vector argument.

Combining Equations 4 and 5, we can rewrite the likelihood of the spectral radiance vector $p(\mathbf{I}(u)$ as

$$p(\mathbf{I}(u)|\Omega) \approx \sum_{m=1}^{M} \frac{\alpha_m}{n_m h} \sum_{v \in \Omega_m} K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|), \tag{6}$$

where the kernel function penalises large arguments, *i.e.* it attains a peak when the argument value is near zero and vanishes when the latter exceeds the kernel width. Examples of kernel functions includes the Gaussian kernel function, which is expressed as

$$K_h^G(q) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{q^2}{2h^2}\right) \tag{7}$$

and the Epanechnikov kernel [32]

$$K_h^E(q) = \begin{cases} \frac{3}{4}\left(1 - \frac{x^2}{h^2}\right) & \text{if } |q| \leq h \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Further, we enforce an additional constraint that imposes the smooth spatial transition of the illumination in the scene. This is because it is often the case that the illumination changes smoothly rather than arbitrarily across the spatial domain of naturally occurring scenes. This implies that the contribution of the constituting illuminants at each pixel varies smoothly. As a result, we introduce a smoothness constraint on the posterior probability $p(L_m|\mathbf{I}(u))$ over the spatial domain $\mathcal{I}$. Such a constraint can be formulated as minimising the gradient norm of the posterior probability as given by

$$\mathcal{R} = \sum_{u \in \mathcal{I}} \sum_{m=1}^{M} \left(\frac{\partial p(L_m|u)}{\partial x}\right)^2 + \left(\frac{\partial p(L_m|u)}{\partial y}\right)^2, \tag{9}$$

where x and y are the coordinates of the pixel u.

Equation 9 states the general smoothness constraint on the continuous spatial domain. To express this constraint over the image lattice, we consider the 4-connected neighbourhood $\mathcal{N}(u)$ around the pixel

u. We then discretise the above equation in terms of the difference of the posterior probability $p(L_m|\cdot)$ between pixel u and a neighbour v, which yields

$$\mathcal{R}^* = \sum_{m=1}^{M} \sum_{u \in \mathcal{I}} \sum_{v \in \mathcal{N}(u)} \left[p(L_m|u) - p(L_m|v) \right]^2, \tag{10}$$

Therefore, the illuminant recovery and segmentation problem is given by a maximisation of the log likelihood of the image radiance spectra subject to the minimisation of the spatial variation of illuminant contributions. Thus, we formulate the objective function as a combination of the log likelihood, *i.e.* the data error term, and the spatial variation of the posterior probability of illuminants, *i.e.* the regularisation term. Hence, our problem amounts to finding the mixing coefficient α_m and the (maximal) sample set of pixels Ω_m for the m^{th} illuminant so as to maximise the function

$$F \quad = \quad \sum_{u \in \mathcal{I}} \log p(\mathbf{I}(u)|\Omega) - \beta \mathcal{R}^*, \tag{11}$$

where $\beta \geq 0$ is the Lagrangian multiplier of the regulariser $\mathcal{R}^*$.

In the following section, we present an iterative approach aimed at maximising the cost function in Equation 11. At each iteration, we re-estimate the weight α_m's and expand the subset of pixels Ω_m so as to maximise the cost function above. We do this making use of the data term by considering the differences between the radiance spectra. Meanwhile, the spatial smoothness constraint propagates illuminant labels to neighbouring pixels. In Algorithm 1 we show the steps of our method in the form of pseudocode. Herein, we have opted to follow the natural progression of the algorithm, first describing the initialisation of the pixel sets $\Omega_1, \Omega_2, \ldots, \Omega_M$. We then turn our attention to the optimisation process and our convergence criterion.

B. Initialisation

As mentioned earlier, we optimise the cost function in Equation 11 using a coordinate descent approach [33]. This comprises optimising the target function with respect to the illuminant cluster parameters in an iterative manner. Here, we model the distribution of pixel radiance spectra illuminated by multiple lights as a mixture, where each image pixel is a combination of the mixture components. In this section, we aim to select initial samples from each of these component (clusters) corresponding to each of the illuminants in the scene. The initialisation procedure is derived from the component separation method described by Kannan *et al.* [21]. To apply this method, we make a weak assumption that the radiance

Algorithm 1 Algorithm to recover the coefficients α, the pixel subsets Ω_m and the probability $p(L_m|u)$.

Require: $\mathbf{I}(u)$: the radiance spectra at each image pixel u.

1: Set M: The number of mixture components.

2: Initialise $\alpha_m^0, \Omega_m^{(0)}, m = 1, \ldots, M$ as in Algorithm 2.

3: **repeat**

4: Compute the mixture coefficient α, recover the posterior probability $p(L_m|u)$ and the update the pixel subset Ω_m using Algorithm 3.

5: **until** Convergence is reached

6: **return** $\alpha_m^{(t)}, \Omega_m^{(t)}, p^{(t)}(L_m|u)$, $\forall m = 1, \ldots, M$.

spectra at pixels illuminated by the same light follow a log-concave distribution. In fact, log-concave densities cover a wide range of common distributions such as Gaussians, the exponential family and the uniform measure over a convex set. Therefore, this assumption does not overly restricts the applicability of our method to real-world scenes.

The algorithm proposed by Kannan *et al.* [21] hinges in the theorem that states that for any mixture of log-concave distributions, the Singular Value Decomposition (SVD) subspace of sample points is close to the means of the samples from the mixture components, where "close" is defined in terms of the sample variances. Further, the method hinges on the intuition that the distance of a point from the mean of its component distribution can only shrink upon projection, thus magnifying the ratio of inter-component to intra-component distances.

Following this intuition, the method in [21] employs a random projection of the input points onto a lower dimensional subspace of the original space. Within this subspace, a projected point with the maximal variance (widest spread) among its neighbourhood along an arbitrary direction in W is deemed to belong to the sparsest domain in its distribution. Using this extremal point as a reference, one can initially select other representatives of the same component by a distance criterion. Specifically, this criterion involves thresholding the ratio of the distance of the input points to the extremal point to the standard deviation of the neighbourhood in the subspace. In [21], the threshold is related to the number of components and the input sample size.

In Algorithm 2, we summarise the initialisation of our method. Throughout the pseudocode in Algorithm 2 we use the symbol $|.|$ to denote the cardinality of a set, $\|.\|$ is the L^2-norm, $\Pi_X(.)$ as the projection

Algorithm 2 Initialise the pixel clusters $\Omega_1, \Omega_2, \ldots, \Omega_M$.

Require: $\mathbf{I}(u)$: the radiance spectra at each image pixel u.

M: The number of mixture components.

{Initialise the current set of pixels} $\mathcal{J}_1 \leftarrow \mathcal{I}$

for $m = 1, \ldots, M$ **do**

  Randomly select $\mathcal{K}_m \subset \mathcal{J}_m$.

  $W_m \leftarrow span(\{\mathbf{I}(u) | u \in \mathcal{K}_m\})$.

  $\forall u \in \mathcal{J}_m \setminus \mathcal{K}_m$, let $\mathbf{Q}(u) \leftarrow \Pi_{W_m}(\mathbf{I}(u))$.

  for $u \in \mathcal{J}_m \setminus \mathcal{K}_m$ **do**

    Define $\mathcal{T}(u) \leftarrow \{v | \|\mathbf{Q}(u) - \mathbf{Q}(v)\| \leq \epsilon_1\}$ {$\mathcal{T}(u)$ is the neighbourhood of u when projected onto W_m.}

    $\bar{\mathbf{P}} \leftarrow \frac{1}{|\mathcal{T}(u)|} \sum_{v \in \mathcal{T}(u)} \mathbf{Q}(v)$.

    $\mathbf{A} \leftarrow [\mathbf{Q}(v_1) - \bar{\mathbf{P}}, \ldots, \mathbf{Q}(v_{|\mathcal{T}(u)|}) - \bar{\mathbf{P}}]$ where $v_i \in \mathcal{T}(u)$.

    $\sigma(u) \leftarrow$ the largest singular value of $\mathbf{A}$.

  end for

  Find $u_0 \in \mathcal{J}_m \setminus \mathcal{K}_m$ such that $\sigma(u_0)$ is maximal.

  Find a direction $V_m \subset W_m$ which maximises $\mathsf{Var}\left(\Pi_{V_m}(\{\mathbf{Q}(u) | u \in \mathcal{K}_m\})\right)$.

  Let U_m be the subspace of W_m orthogonal to V_m.

  Find the maximal pixel subset $\Omega_m^0 \subset \mathcal{J}_m \setminus \mathcal{K}_m$ such that $\|\mathbf{I}(u) - \mathbf{I}(u_0)\| \leq \epsilon_2$ and $\|\Pi_{U_m}(\bar{\mathbf{I}}(u)) - \Pi_{U_m}(\mathbf{I}(u_0))\| \leq \epsilon_3 \ \forall u \in \mathcal{L}_m$, where $\bar{\mathbf{I}}(u) = \frac{\sum_{u \in \Omega_m} \mathbf{I}(u)}{|\Omega_m|}$.

  $n_m^0 \leftarrow |\Omega_m^0|$.

  $\alpha_m^0 \leftarrow \frac{n_m^0}{\sum_{m=1}^M n_m^0}$.

  Update the remaining set of pixels $\mathcal{J}_{m+1} \leftarrow \mathcal{J}_m \setminus \Omega_m^0$.

end for

return $\Omega_m^0, \forall m = 1, \ldots, M$.

operator onto a subspace X, $\mathsf{Var}(.)$ is the variance operator and ϵ_1, ϵ_2 and ϵ_3 are constants.

The initialisation takes place over M iterations, in each of which we select a subset of sample pixels Ω_m that are representative of the component distribution corresponding to the m^{th} light. At the beginning of each iteration, we randomly sample a subset of pixels $\mathcal{K}_m$ from the unassigned pixels $\mathcal{J}_m$ in the image,

i.e. those pixels that have not been sampled at previous iterations, as shown in Line 3. Subsequently, we obtain the projection $\mathbf{Q}(u)$ onto the M-dimensional subspace W_m spanned by the radiance spectra at the pixels in $\mathcal{K}_m$, whose basis vectors are found via SVD, as shown in Line 5. At this stage, the random subset $\mathcal{K}_m$ used to obtained the low-dimensional subspace is discarded from $\mathcal{J}_m$ to ensure that the samples selected for each component is independent from W_m. Later, as shown between Lines 7 and 10, we find the largest variance $\sigma(u)$ of the projected neighbours of $\mathbf{Q}(u)$ in the subspace W_m. This is achieved by first constructing a matrix $\mathbf{A}$ whose columns are the mean-centred projection of the neighbours of $\mathbf{Q}(u)$ in W_m. The largest variance $\sigma(u)$ is found to be the largest singular value of $\mathbf{A}$.

In Line 12, the algorithm aims to find a pixel $u_0 \in \mathcal{J}_m \backslash \mathcal{K}_m$ with the maximal variance $\sigma(u_0)$. Thereafter, between Lines 13 and 14, we find a subspace $U_m \subset W_m$ that is orthogonal to the direction $V_m \subset W_m$ along which the variance of the projected point is maximal. At the end of each iteration, we choose the initial sample pixels Ω_m^0 under the m^{th} illuminant as those close to both the original radiance space and the projection subspace U_m, as shown in Line 15.

C. Optimisation

We summarise all the steps of our optimisation approach in Algorithm 3. Note that our optimisation yields at output the weight α_m, the maximal pixel subsets Ω_m for each of the illuminants in the scene and the abundance of each light L_m at each pixel u, *i.e.* $p(L_m|u)$.

To recover the mixed illumination at each pixel, we compute the power spectrum of the illuminant L_m from the subset of pixels Ω_m. This is achieved by applying a colour constancy method, such as the Grey-World [7], Grey-Edge [10] and White-Patch algorithm [6] to this subset of pixels. We can then re-estimate the per-pixel illumination using the probabilities $p(L_m|u)$, as follows

$$L(u) = \sum_{m=1}^{M} p(L_m|u) L_m. \tag{12}$$

For our optimisation, we depart from the initial set of pixels and maximise the cost function in Equation 11. Using the approximation in Equation 6, we can rewrite this cost function as

$$F \approx \sum_{u \in \mathcal{I}} \log \left(\sum_{m=1}^{M} \frac{\alpha_m}{n_m} \sum_{v \in \Omega_m} K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|) \right) - \beta \mathcal{R}^* - |\mathcal{I}| \log h, \tag{13}$$

where $|\mathcal{I}|$ is the total number of image pixels.

Since the last term on the right hand side of Equation 13 is a constant, we can simply remove it from further consideration. We also note that the current setting assumes an implicit relationship between the

Algorithm 3 Compute the mixture coefficient α, recover the posterior probability $p(L_m|u)$and the update the pixel subset Ω_m.

Require: $\mathbf{I}(u)$: the radiance spectra at each image pixel u.

M: The number of mixture components.

Initialise $\alpha_m^0, \Omega_m^{(0)}, m = 1, \ldots, M$ as in Algorithm 2.

repeat

Update $p^{(t+1)}(L_m|u)$ by solving Equation 16.

$\alpha_m^{(t+1)} = \frac{\sum_{u\in\mathcal{I}} p^{(t+1)}(L_m|u)}{|\mathcal{I}|}$ (Equation 18).

$\forall u \in \mathcal{I}$, define $l(u) = \text{argmax}_{1\leq m\leq M}\, p^{(t+1)}(L_m|u)$.

for $m = 1, \ldots, M$ **do**

Define $\Upsilon_m = \{u \in \mathcal{I} | l(u) = m \text{ and } p^{(t+1)}(L_{l(u)}|u) > t_l\}$.

$\forall u \in \Upsilon_m$, let $\mathcal{S}^r(u)$ be the set of r pixels with $\mathbf{I}(v)$ nearest to $\mathbf{I}(u)$.

$\Omega_m^{(t+1)} \leftarrow \Omega_m^{(t)} \cup (\cup_{u\in\Upsilon_m} \mathcal{S}^r(u))$.

end for

$t \leftarrow t + 1$.

until $\alpha_m^{(t)}, \Omega_m^{(t)}, p^{(t)}(L_m|u)$ are close to $\alpha_m^{(t-1)}, \Omega_m^{(t-1)}, p^{(t-1)}(L_m|u)$ within a tolerance.

return $\alpha_m^{(t)}, \Omega_m^{(t)}, p^{(t)}(L_m|u)$, $\forall m = 1, \ldots, M$.

posterior probability $p(L_m|u)$ and the likelihood of the image radiance $p_m(u)$ as given by Bayes rule as follows

$$p(L_m|u) = \frac{\alpha_m p_m(u)}{\sum_{q=1}^{M} \alpha_q p_q(u)}. \tag{14}$$

As a result, we can constrain the optimisation using Equation 14. By adopting a coordinate-ascent optimisation approach. Our iterative algorithm consists of two interleaving steps in each iteration. The first step re-estimates the posterior probability distribution $p(L_m|\mathbf{I}(u))$, whereas the second one estimates the mixing coefficients α and expands the set Ω_m for each illuminant once the posterior probability distribution is at hand. In the following subsections, we index these variables to the iteration number t and describe these steps in detail.

1) Updating the posterior probabilities $p(L_m|u)$*:* With the initial mixture coefficients and sample pixels under each illuminant, we proceed to estimate the posterior probability. Let us use the shorthand $p^{(t)}(L_m|u)$ to denote the estimate of $p(L_m|u)$ given the values of α and the sample pixels set Ω at the t^{th} iteration.

Also, let $\bar{p}^{(t)}(L_m|u)$ denote the Bayessian estimate of $p(L_m|u)$ given by Equation 14 with respect to $\alpha^{(t)}$ and $p_m^{(t)}(u) = p(\mathbf{I}(u)|\Omega^{(t)})$.

Since only the second term, *i.e.* the spatial smoothness term, of the cost function in Equation 13 is dependent on $p(L_m|u)$, this optimisation step can be viewed as that aiming to find a posterior probability which minimises

$$F_1 = \beta \sum_{m=1}^{M} \sum_{u \in \mathcal{I}} \sum_{w \in \mathcal{N}(u)} \left[p(L_m|u) - p(L_m|w) \right]^2 + \gamma \sum_{m=1}^{M} \sum_{u \in \mathcal{I}} \left(p(L_m|u) - \bar{p}(t)(L_m|u) \right)^2, \tag{15}$$

where $\gamma \geq 0$ is the weight of the Bayesian constraint.

Setting the derivative of F_1 with respect to $p(L_m|w)$ to zero, we obtain

$$(\beta |N(u)| + \gamma) p(L_m|u) - \beta \sum_{w \in \mathcal{N}(u)} p(L_m|w) = \gamma \bar{p}(t)(L_m|u) \tag{16}$$

Note that, following Equation 16, there exists a linear system of equations with respect to the variables $\{p(L_m|u)|u \in \mathcal{I}\}$ for each illuminant m. To solve such a system under the constraints that $0 \leq p(L_m|u) \leq 1$ and $\sum_{m=1}^{M} p(L_m|u) = 1$, we use the standard linear solver technique in [34].

2) Updating the mixing coefficients α_m: Having estimated the posterior probability, we now estimate the mixture coefficients α_m. We note that only the first term of the function in Equation 13, which we denote τ_1, is dependent on these two subsets of variables. For the current estimate of the posterior distribution $p^{(t+1)}(L_m|u)$, the expected value of this term is

$$\mathbb{E}(\tau_1) = \sum_{u \in \mathcal{I}} \sum_{m=1}^{M} \log \alpha_m p^{(t+1)}(L_m|u) + \sum_{u \in \mathcal{I}} \sum_{m=1}^{M} \log \left(\frac{\sum_{v \in \Omega_m} K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|)}{n_m} \right) p^{(t+1)}(L_m|u). \tag{17}$$

Thus, we maximise the function above with respect to α under the constraint $\sum_{m=1}^{M} \alpha_m = 1$. By setting the derivative of the equation above with respect to α to zero, we obtain

$$\alpha_m^{(t+1)} = \frac{\sum_{u \in \mathcal{I}} p^{(t+1)}(L_m|u)}{|\mathcal{I}|} \tag{18}$$

3) Expanding the subsets of pixels Ω_m: Now we turn our attention to expanding the subset of sample pixels Ω_m for each illuminant L_m. This implies that, for each m, we find a subset Ω_m that maximises the weighted sum $\sum_{u \in \mathcal{I}} \log \left(\frac{\sum_{v \in \Omega_m} K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|)}{n_m} \right) p^{(t+1)}(L_m|u)$, where the weights $p^{(t+1)}(L_m|u)$ are obtained in the previous step.

Before proceeding further, we comment briefly on our notation. For each pixel u in the image, we define its dominant light $l(u)$ to be the one with the largest posterior probability, *i.e.* such that $l(u) =$

$\text{argmax}_{1\leq m\leq M}\, p(L_m|u)$. Also, we denote Υ_m as the subset of pixels with the m^{th} illuminant as their dominant illuminant $\Upsilon_m = \{u \in \mathcal{I} | l(u) = m\}$.

To expand Ω_m we adopt a greedy approach. We note that the kernel value $K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|)$ is large when the vectors $\mathbf{I}(u)$ and $\mathbf{I}(v)$ are close. For each pixel $u \in \Upsilon_m$, to maximise $K_h(\|\mathbf{I}(u) - \mathbf{I}(v)\|)$, we choose a subset $\mathcal{S}^r(u)$ of r-nearest neighbours of u in the spectral radiance space, where r is a preset constant. The union of the subsets $\mathcal{S}^r(u)$, where $u \in \Upsilon_m$, are then augmented to the current subset of pixels under the m^{th} illuminant. In other words, the sample pixels under the m^{th} illuminant is updated using the rule $\Omega_m^{(t+1)} \leftarrow \Omega_m^{(t)} \cup (\cup_{u\in\Upsilon_m} \mathcal{S}^r(u))$. The algorithm terminates when the estimated mixing coefficients and the subsets of sample pixels $\Omega_m, m = 1, \ldots, M$ converge, *i.e.* no further updates are made.

IV. Experiments

In this section, we illustrate the utility of our method for the segmentation and estimation of non-uniform illumination from a single image. In our experiments, we have compared our method to a number of alternatives. These are the method of Gijsenij *et al.* [19], the Retinex algorithm [35] and the approach proposed by Ebner [15]. For our method and Gijsenij *et al.*'s, we apply a variety of existing colour constancy methods, such as the Grey World [7], Grey Edge [10] and White Patch [6] to recover the illuminant colour in each illuminant segment. The comparison presented here involves several variants of these, including the general Grey World (general GW) [19] with L^8-norm, the first-order Grey Edge (1^{st} GE) and 2^{nd}-order Grey Edge (2^{nd} GE). Gijsenij *et al.*'s method [19] was performed with a grid-based sampling strategy with a patch size of 10×10.

For comparison purposes, we have employed k-means to segment the illumination images produced by Ebner's [15], Retinex [35] and Gijsenij *et al.*'s method [19]. Note that these methods do take the correct number of illuminants in the scene at input. This contrasts with our method, which does not require the number of illuminants to be supplied a priori. Rather, the initialisation step in Algorithm 2 determines the number of illuminants automatically. In addition, for our method, we have employed a Gaussian kernel to estimate the probability density function of the radiance vector $\mathbf{I}(u)$ in Equation 4. Here we set the kernel width h to be the standard deviation of all the pair-wise distances between the radiance spectra in the input scene.

For our experiments, we have used a number of datasets. The first of these consists of hyperspectral images of eight rural and urban scenes collected by Foster *et al* [36]. To simulate the spatial variation

Our method	77.09 ± 19.10				
Ebner's [15]	43.02 ± 9.76				
Retinex [35]	44.76 ± 12.78				
	Grey World	White Patch	General GW	1^{st} GE	2^{nd} GE
Gijsenij *et al*'s [19]	75.19 ± 22.05	76.81 ± 19.74	74.35 ± 21.02	73.78 ± 19.73	74.00 ± 20.20

TABLE I

THE SEGMENTATION ACCURACY ϵ (IN $\%$) FOR THE HYPERSPECTRAL IMAGE DATASET *Foster-spectral*.

Our method	65.77 ± 16.12				
Ebner's [15]	47.07 ± 7.00				
Retinex [35]	48.21 ± 6.73				
	Grey World	White Patch	General GW	1^{st} GE	2^{nd} GE
Gijsenij *et al.*'s [19]	70.97 ± 20.97	71.20 ± 19.81	70.89 ± 20.34	67.95 ± 19.75	69.24 ± 19.91

TABLE II

THE SEGMENTATION ACCURACY ϵ (IN $\%$) FOR THE PSEUDOCOLOUR IMAGE DATASET *Foster-colour* WHICH WE SYNTHESIZE FROM THE DATASET COLLECTED BY FOSTER *et al.* [36].

of illumination, we divide each image into quadrants and apply the illumination spectrum of different light to each region. This procedure aims to apply two to three different illuminant spectra across the scene in a similar fashion to the method reported in [19]. These power spectra act as filters which are randomly selected from a set of 81 illuminant spectra collected by Barnard *et al.* [37]. We also generate the pseudocolour counterparts from these images with the standard colour matching functions [38]. We denote the hyperspectral dataset *Foster-spectral* and the corresponding pseudocolour version *Foster-colour*.

In addition, we perform experiments on two other datasets of colour images illuminated by multiple illuminants. One of them consists of the natural scenes captured outdoors by Gijsenij *et al.* [19]. The

Our method	77.69 ± 19.41				
Ebner's [15]	61.24 ± 18.00				
Retinex [35]	72.16 ± 13.33				
	Grey World	White Patch	General GW	1^{st} GE	2^{nd} GE
Gijsenij *et al*'s [19]	71.13 ± 25.79	71.71 ± 21.09	69.47 ± 24.93	61.08 ± 21.52	63.99 ± 23.75

TABLE III

THE SEGMENTATION ACCURACY ϵ (IN $\%$) FOR THE NATURAL IMAGES CAPTURED BY GIJSENIJ *et al.* [19].

ground truth colour of the light sources were obtained by placing several gray surfaces over the scene and manually annotating the area for every light source. We denote this dataset herein *Gijsenij*. The other dataset of colour images is that acquired by Bleier *et al.* [4] [1]. The dataset contains 408 images of four scenes acquired under two or three illuminants. It also includes ground truth illumination colour images acquired using same scenes spray-painted in gray. We have denoted this dataset after its author as *Bleier*.

A. Illumination Region Segmentation

We first examine the accuracy of the segmented illuminant regions resulting from the above methods. To provide a quantitative evaluation of the segmentation results, we adopt an accuracy measure based on the overlap between the ground-truth and the estimated region for each illuminant. This measure has been employed elsewhere to quantify the segmentation accuracy of the PASCAL Visual Object Classes (VOC) Challenge [39]. Let A_m^{gt} and A_m^{est} be respectively the ground-truth and the estimated region of the m^{th} illuminant. The segmentation error is then formulated as

$$\epsilon = \frac{1}{M}\sum_{m=1}^{M}\frac{A_m^{gt}\bigcap A_m^{est}}{A_m^{gt}\bigcup A_m^{est}}, \tag{19}$$

where $A_m^{gt}\bigcap A_m^{est}$ denotes the intersection of the estimated and ground truth regions of the m^{th} light and $A_m^{gt}\bigcup A_m^{est}$ their union.

In Figure 2, we provide a qualitative demonstration of the illuminant segmentation results. The figure includes sample illuminant segmentation maps for three scenes, one from each dataset, with the segmentation maps for each method shown in separate columns. The first column shows the input image while the remaining ones show the illuminant regions recovered by the different methods. In the figure, each colour in the segmentation maps corresponds to a distinct illuminant label. In the last column, we show the results produced by the variant of Gijsenij *et al.*'s method [19] that employs the Grey-World assumption for local illumination estimation. This is the variant implemented by Gijsenij *et al.* [19] that generally yields the highest performance and, as we will show later in Tables I, II, and III, this is consistent across our experiments. As observed from this figure, our method delivers more accurate segmentation maps than the others. While Ebner's and the Retinex algorithm often result in small discontinuous fragments, Gijsenij *et al.*'s method [19] yields patchy and isolated regions due to the nature of its grid-based sampling strategy.

[1]The dataset is accessible at http://www5.informatik.uni-erlangen.de/research/data/multi-illuminant-dataset/

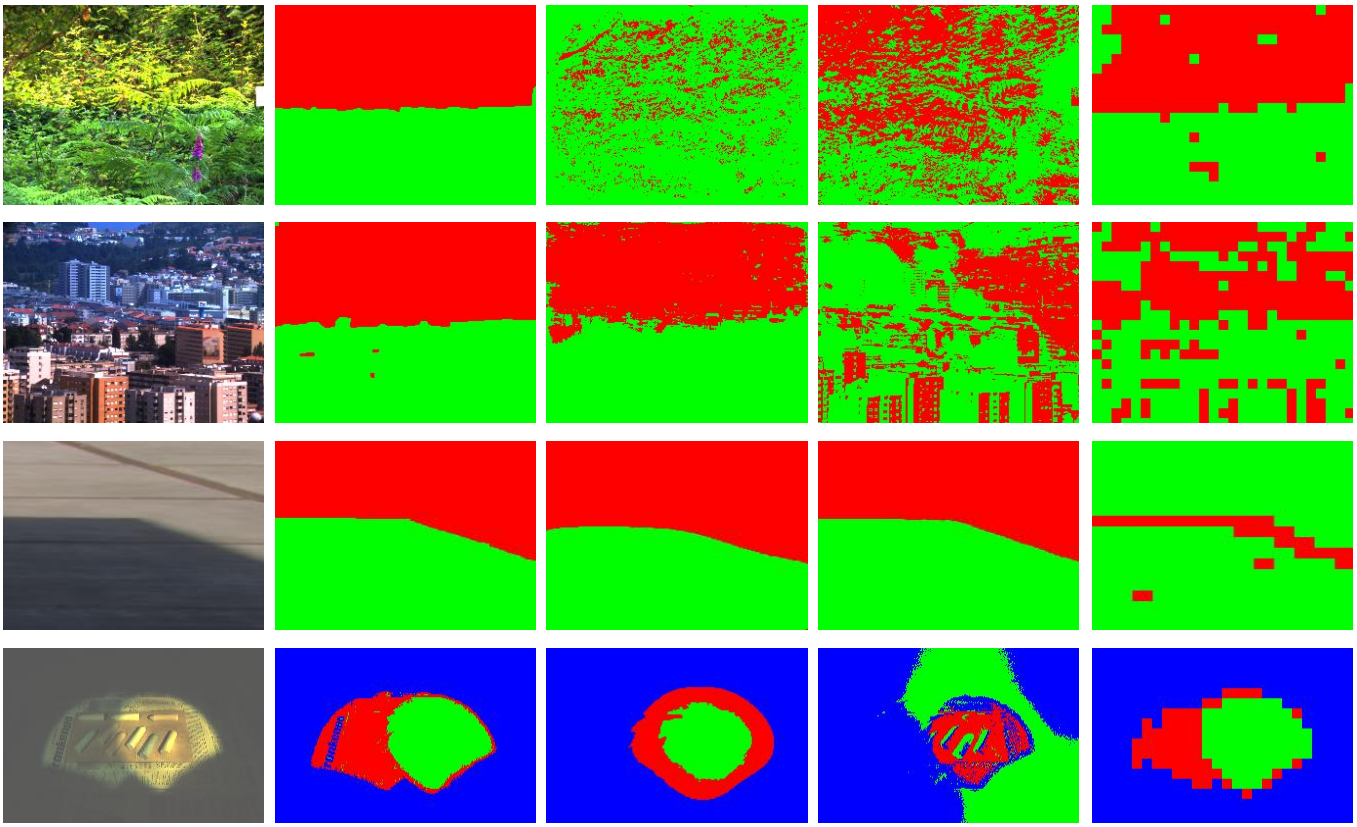


Fig. 2. Illuminant regions segmenteds by our method and a number of alternatives. In each row (from top to bottom), we show results for a sample image from each of our datasets, *i.e. Foster-spectral*, *Foster-colour*, *Gijsenij* and *Bleier*. First column: Input image; Second to last columns: Illuminant segments yielded by our method, Ebner's [15], Retinex method [35] and Gijsenij *et al.*'s [19], respectively.

In Table I, we show the segmentation accuracy as defined in Equation 19 for the spectral images in the dataset *Foster-spectral*. Similarly, Table II shows the segmentation results for the pseudocolour version of these images, *i.e.* the *Foster-colour* dataset. For the former dataset, our method slightly outperforms the variants of Gijsenij *et al.*'s method [19] and greatly outperforms the others. In addition, both of these methods outperform the remaining alternatives for the two datasets. Further, we show in Table III the illumination segmentation results for the natural images captured by Gijsenij *et al.* [19]. For this dataset, our approach outperforms all the alternatives by at least 5.5%.

B. Illumination Spectrum Estimation

Now, we turn our attention to the efficacy of our method in estimating the illuminants in the scene. To quantify the accuracy of the estimated illuminants for each image, we compute the Euclidean angle between the spatial average of the ground truth and the estimated illuminant colour vectors. In Tables IV

	Grey-World	White-Patch	General GW	1^{st} GE	2^{nd} GE
Our method	25.03	19.23	20.16	20.59	20.35
Gijsenij *et al.* 's [19]	25.87	21.30	23.98	21.80	21.22
Ebner's [15]	38.04				
Retinex [35]	28.44				

TABLE IV

MEDIAN ANGULAR ERRORS (IN DEGREES) FOR THE ILLUMINANT SPECTRUM ESTIMATED FROM THE HYPERSPECTRAL IMAGE DATASET *Foster-spectral* [36].

	Grey-World	White-Patch	General GW	1^{st} GE	2^{nd} GE
Our method	14.76	13.60	15.28	16.26	15.77
Gijsenij *et al.* 's [19]	12.69	12.59	13.06	16.58	15.77
Ebner's [15]	21.46				
Retinex [35]	18.44				

TABLE V

MEDIAN ANGULAR ERRORS (IN DEGREES) FOR THE ILLUMINANT COLOUR IN THE PSEUDOCOLOUR IMAGE DATASET *Foster-colour*, WHICH WAS SYNTHESIZED FROM THE DATASET COLLECTED BY FOSTER *et al.* [36].

and V, we report the median angular error for the illuminant colours estimated from the hyperspectral image dataset *Foster-spectral* and its pseudocolour version *Foster-colour*, respectively. Similarly, we report the angular error for the *Gijsenij* dataset in Table VI and those for the *Bleier* dataset in Table VII.

From the tables, we can conclude that, on the *Foster-spectral* and *Foster-colour* datasets, all the variants of our method are comparable to Gijsenij *et al.*'s method and outperform the others by at least three degrees when performing. The variants of our method and Gijsenij *et al.*'s method that employ White-Patch to estimate the illumination achieve the highest accuracy. This is due to the presence of materials with gray

	Grey-World	White-Patch	General GW	1^{st} GE	2^{nd} GE
Our method	3.45	2.97	2.86	3.25	3.27
Gijsenij *et al.* 's [19]	6.33	6.27	6.26	6.52	5.98
Ebner's [15]	5.82				
Retinex [35]	5.47				

TABLE VI

MEDIAN ANGULAR ERRORS (IN DEGREES) FOR THE ILLUMINANT COLOUR ESTIMATED FROM THE NATURAL IMAGES ACQUIRED BY GIJSENIJ *et al.* [19].

	Grey-World	White-Patch	General GW	1^{st} GE	2^{nd} GE
Our method	1.18	1.77	1.56	3.39	3.32
Gijsenij *et al.* 's [19]	4.71	6.12	4.78	14.89	13.87
Ebner's [15]	4.76				
Retinex [35]	5.18				

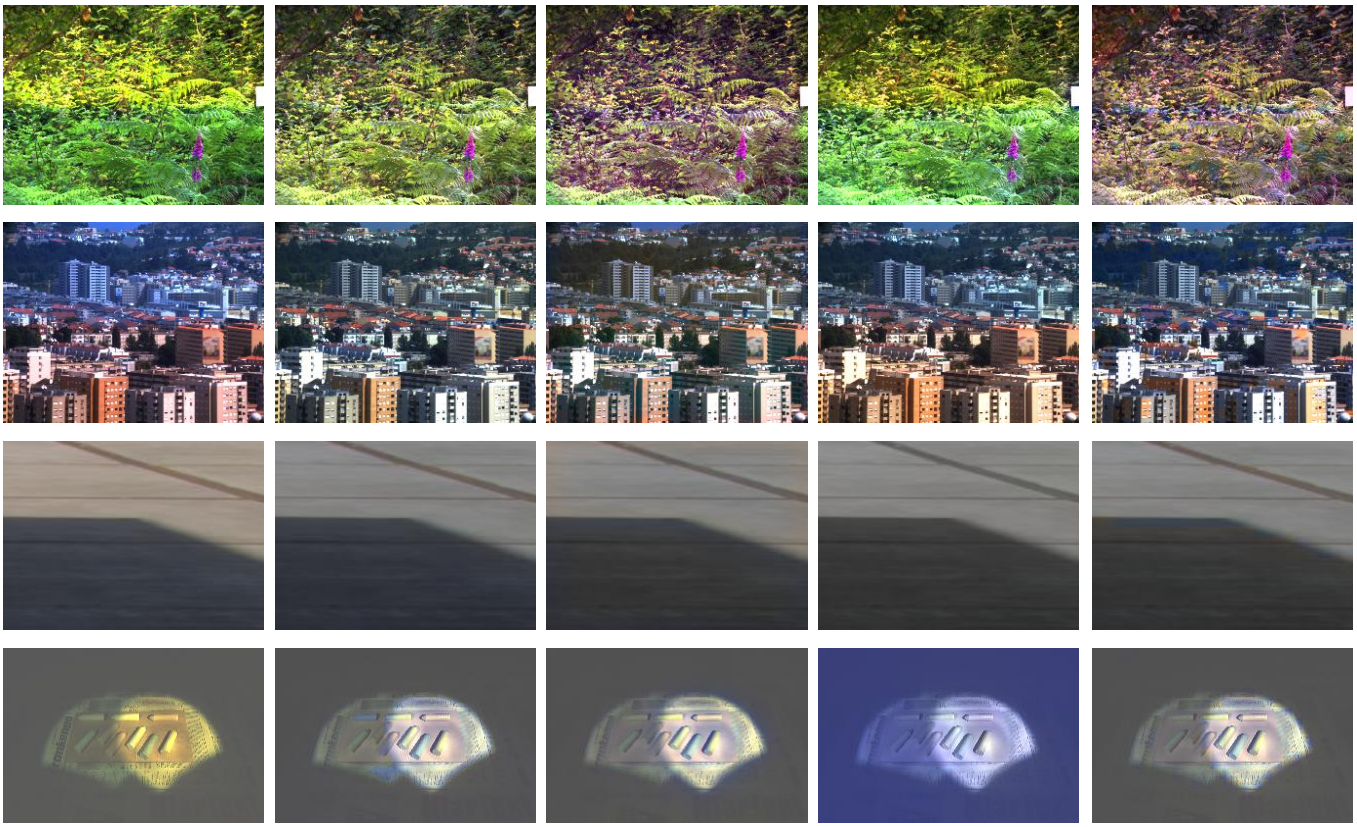
TABLE VII

MEDIAN ANGULAR ERRORS (IN DEGREES) FOR THE ILLUMINANT COLOUR ESTIMATED FROM THE DATASET REPORTED IN BLEIER *et al.* [4].

or white reflectance in the scene. However, the estimation accuracy does not appear to vary greatly with respect to the colour constancy method used in conjunction with our approach and Gijsenij *et al.*'s one.

On the *Gijsenij* and *Bleier* datasets, all the variants of our method significantly outperform the remaining ones, achieving an estimation error no more than half of those produced by the alternatives. This good performance is justified by the high accuracy of our method in locating illuminant regions. Gijsenij *et al.*'s method [19] does not perform as well as ours due to its assumption of constant illuminant on every sampled patch. Ebner's local averaging method does not seem to perform well on these datasets since each image only consists of a small number of materials, which does not satisfy the Grey-World assumption. Lastly, the Retinex implementation in [35] relies on an initialisation based on the White-Patch assumption, which does not work for these images since they lack white materials.

Finally, we show colour correction results based on the spatially varying illuminant colour estimated by each of the methods under consideration. To perform colour correction on the original images, we have substituted the illuminant per pixel with a single white light as the reference illuminant. In Figure 3, we show the colour-balanced output for the methods above in separate columns. Note that the colour-corrected images using the illuminant colour estimated by our method exhibit a good level of chromaticity consistency over the scene, especially within the same material regions across illuminant boundaries. Furthermore, our method yields the most neutral colour corrected imagery. In contrast, the alternatives often exhibit colour shades strongly influenced by the original illuminant and colour bleeding along illuminant boundaries.



Input image Our method Ebner's [15] Retinex [35] Gijsenij *et al.*'s [19]

Fig. 3. Colour correction with respect to a white light source. In each row (from top to bottom), we show results for a sample image from each of the following datasets, *i.e. Foster-spectral*, *Foster-colour*, *Gijsenij* and *Bleier*. First column: Input images. Remaining columns, left-to-right: Colour corrected images using the illuminant recovered by the Grey-World variant of our method, Ebner's [15], Retinex method [35] and the Grey-World variant of Gijsenij *et al.*'s [19], respectively.

V. Conclusions

In this paper, we have presented an algorithm to segment the illuminant regions and to estimate the illumination power spectrum from a single image of a scene lit by multiple light sources. Our method exploits the probability distribution of the spectral radiance at pixels illuminated by the same illuminant to cast the problem as an estimation of mixture components in the spectral radiance space. Based on this approach, the problem becomes that of maximising the likelihood of the image pixels being illuminated by the scene illuminants while enforcing the smooth variation of the light contribution over the spatial domain. To this end, we use a kernel density estimator as applied to a subset of pixels illuminated by each illuminant. We tackle the illumination segmentation problem using a coordinate ascent scheme, which involves update operations on the mixture coefficients and the posterior probability of each light per pixel. Once the pixel clusters are in hand, we estimate the illuminant spectrum for using standard colour

constancy methods for single illuminants, and re-estimate the illumination per pixel with the posterior probability of each light. We have also shown results on a number of image datasets and compared to alternatives elsewhere in the literature. Our experiments show the utility of our method for the illumination region segmentation, illumination colour estimation and colour correction.

ACKNOWLEDGMENT

NICTA is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

REFERENCES

[1] T. Zickler, S. P. Mallick, D. J. Kriegman, and P. N. Belhumeur, "Color subspaces as photometric invariants," *International Journal of Computer Vision*, vol. 79, no. 1, pp. 13–30, 2008.

[2] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, "Light mixture estimation for spatially varying white balance," ser. SIGGRAPH '08, 2008, pp. 70:1–70:7. [Online]. Available: http://doi.acm.org/10.1145/1399504.1360669

[3] J. Lopez-Moreno, S. Hadap, E. Reinhard, and D. Gutierrez, "Compositing images through light source detection," *Computer & Graphics*, vol. 34, no. 6, pp. 698–707, Dec. 2010. [Online]. Available: http://dx.doi.org/10.1016/j.cag.2010.08.004

[4] M. Bleier, C. Riess, S. Beigpour, E. Eibenberger, E. Angelopoulou, T. Tröger, and A. Kaup, "Color constancy and non-uniform illumination: Can existing algorithms work?" in *ICCV Workshops*, 2011, pp. 774–781.

[5] C. P. Huynh and A. Robles-Kelly, "A solution of the dichromatic model for multispectral photometric invariance," *International Journal of Computer Vision*, vol. 90, no. 1, pp. 1–27, 2010.

[6] E. H. Land and J. J. Mccann, "Lightness and retinex theory," *J. Opt. Soc. Am*, vol. 61, pp. 1–11, 1971.

[7] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, no. 1, pp. 337–350, 1980.

[8] J. J. McCann, J. A. Hall, and E. H. Land, "Color mondrian experiments: the study of average spectral distributions," *Journal of Optical Society America*, vol. 67, p. 1380, 1977.

[9] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Color Imaging Conference*, 2004, pp. 37–41.

[10] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2207–2214, 2007.

[11] A. Bousseau, S. Paris, and F. Durand, "User-assisted intrinsic images," ser. SIGGRAPH Asia '09, 2009, pp. 130:1–130:10. [Online]. Available: http://doi.acm.org/10.1145/1661412.1618476

[12] I. Boyadzhiev, K. Bala, S. Paris, and F. Durand, "User-guided white balance for mixed lighting conditions," *ACM Transactions on Graphics*, vol. 31, no. 6, pp. 200:1–200:10, Nov 2012. [Online]. Available: http://doi.acm.org/10.1145/2366145.2366219

[13] M. Ebner, *Color constancy*, ser. Imaging Science and Technology. Wiley- IS&T, 2007.

[14] G. D. Finlayson, B. V. Funt, and K. Barnard, "Color constancy under varying illumination," in *ICCV*, 1995, pp. 720–725.

[15] M. Ebner, "Color constancy using local color shifts," in *European Conference on Computer Vision*, 2004, pp. 276–287.

[16] Y. Wang and D. Samaras, "Estimation of multiple illuminants from a single image of arbitrary known geometry," in *European Conference on Computer Vision*, 2002, pp. 272–288.

[17] G. D. Finlayson, C. Fredembach, and M. S. Drew, "Detecting illumination in images," in *IEEE International Conference on Computer Vision*, 2007.

[18] K. Barnard, G. D. Finlayson, and B. V. Funt, "Color constancy for scenes with varying illumination," *Computer Vision and Image Understanding*, vol. 65, no. 2, pp. 311–321, 1997.

[19] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 697–707, 2012.

[20] C. Riess, E. Eibenberger, and E. Angelopoulou, "Illuminant Color Estimation for Mixed-Illuminant Real-World Scenes," in *ICCV Workshops*, 2011, pp. 774–781.

[21] R. Kannan, H. Salmasian, and S. Vempala, "The spectral method for general mixture models," in *18th Annual Conference on Learning Theory (COLT*, 2005, pp. 444–457.

[22] E. Parzen, "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[23] A. Gijsenij, R. Lu, and T. Gevers, "Color constancy for multiple light sources." *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 697–707, 2012.

[24] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," *International Journal of Computer Vision*, vol. 52, no. 1, pp. 7–23, 2003.

[25] J. Cohen, "Dependency of the spectral reflectance curves of the munsell color chips," *Psychonomic Science*, vol. 1, no. 12, pp. 369–370, Oct 1964.

[26] L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," *Journal of Optical Society America A*, vol. 3, no. 10, pp. 1673–1683, Oct 1986. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-3-10-1673

[27] D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra," *J. OPT. SOC. AM. A*, vol. 9, pp. 1905–1913, 1992.

[28] D. L. Ruderman, T. W. Cronin, and C.-C. Chiao, "Statistics of cone responses to natural images: implications for visual coding," *Journal of Optical Society America A*, vol. 15, no. 8, pp. 2036–2045, August 1998. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-15-8-2036

[29] T.-W. Lee, T. Wachtler, and T. J. Sejnowski, "The spectral independent components of natural scenes," in *Proceedings of the First IEEE International Workshop on Biologically Motivated Computer Vision*, ser. BMVC '00. London, UK, UK: Springer-Verlag, 2000, pp. 527–534. [Online]. Available: http://dl.acm.org/citation.cfm?id=648247.761413

[30] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized Gamut Mapping using Image Derivative Structures for Color Constancy," *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 127–139, 2010.

[31] J. Hoffbeck and D. Landgrebe, "A method for estimating the number of components in a normal mixture density function," in *Proceedings of Geoscience and Remote Sensing Symposium*, vol. 4, 2000, pp. 1675–1677.

[32] D. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. Wiley, 1992.

[33] J. Nocedal and S. Wright, *Numerical Optimization*. Springer, 2000.

[34] S. Wright, *Primal-Dual Interior-Point Methods*, ser. Miscellaneous Bks. Society for Industrial and Applied Mathematics, 1997.

[35] B. V. Funt, F. Ciurea, and J. J. McCann, "Retinex in MATLAB^{TM}," *J. Electronic Imaging*, vol. 13, no. 1, pp. 48–57, 2004.

[36] D. H. Foster, S. M. Nascimento, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual Neuroscience*, vol. 21, pp. 331–336, 4 2004. [Online]. Available: http://journals.cambridge.org.virtual.anu.edu.au/article_S0952523804213335

[37] K. Barnard, L. Martin, B. Funt, and A. Coath, "A Data Set for Colour Research," *Color Research and Application*, vol. 27, no. 3, pp. 147–151, 2002. [Online]. Available: http://www.cs.sfu.ca/~colour/data

[38] W. S. Stiles and J. M. Burch, "Interim report to the Commission Internationale de l'Éclairage Zurich, 1955, on the National Physical Laboratory's investigation of colour-matching," *Optica Acta*, vol. 2, pp. 168–181, 1955.

[39] M. Everingham, L. Gool, C. K. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010. [Online]. Available: http://dx.doi.org/10.1007/s11263-009-0275-4