https://doi.org/10.2352/ISSN.2470-1173.2017.18.COLOR-060

Illumination and Reflectance Spectra Separation of Hyperspectral Image Data under Multiple Illumination Conditions

Xiaochuan Chen[1] *, Mark S. Drew*[1] *, Ze-Nian Li*[1] *;* [1] *Simon Fraser University,* {*xca64,mark,li*}*@cs.sfu.ca*

Abstract

Recently, a remarkably simple method was developed to solve the illumination and reflectance spectra separation problem (IRSS) based on the standard low-dimensionality assumption of reflectance. However, because this method assumes the scene is under one uniform illumination, it can not handle scene contains multiple illuminations or dominant shadows. In this paper, we address this problem by formulating the multiple illuminations and reflectance separation problem as a Conditional Random Field (CRF) optimization task over local separations. We then improve local illumination and reflectance separation by incorporating spatial information in each local patch.

1. Introduction

Spectral distributions in hyperspectral images result from the product of the illumination spectra, surface reflection, and the effects of camera sensors' band sensitivities. Separating illumination and reflectance signals from an observed hyperspectral image has been a longstanding problem in computer vision [3, 4, 7].

In fact the human visual system is capable of separating out the surface reflectance when observing a scene — a psychophysical phenomenon denoted color constancy. The analogous goal in computer science is often to estimate illumination spectra as a first step and then recover the reflectance spectra.

In general, two main streams have been used in solving this problem. The first is the use of hyperspectral image statistics information to estimate illumination and then separate reflectance information from the image. Another stream is to separate illumination and reflectance directly from the observed hyperspectral signal. This process is under-constrained, since the goal consists of twice the number variables as observations. A typical way to resolve this issue is by introducing a low dimensional subspace for reflectance or illumination or both. Remarkably, [17] addresses a more general scenario where surface reflectance is assumed to fall into a low dimensional subspace while no constraint is placed on illumination.

In [17], the separation problem is modelled as a low-rank matrix factorization, with disentangled illumination and reflectance found by solving a small convex quadratic program. Their method is robust even when the scene is under general, non-smooth, illumination. However, nevertheless it can if their exist dominant shadows or multiple illuminations in the scene.

In this paper we re-examine Zheng's method [17] with a view to developing extensions which we find further improve the separation. Moreover, we introduce a CRF (Conditional Random Field) model over the results of separations from method [17] to handle scenes containing multiple illuminations or dominant shadows.

The paper is organized as follows. In sections 2 we discuss prior research that forms the foundation of the current work. In section 3 reviewed Zheng's method [17] for illumination and reflectance separation. In section 4 we propose extensions that improve separation for illumination and reflectance. As well, we introduced a CRF model for segmenting different illuminated patches. In section 5, we show results for the method [17] alongside our separation results, when contending with multiple illuminations in the scene. At final section, we concluded this paper.

2. Related Work

2.1 Illumination and Reflectance Spectra Separation

Most works on illumination and reflectance spectra separation can be summarized into two categories, either separating a single hyperspectral image into illumination and reflectance image, or with an eye to the colour constancy task of estimating illumination for RGB images.

Classically, for this signal separation task, Ho et al. [7] first made use of the low dimension model for both illumination and reflectance. To improve accuracy of the separation, Chang et al. [3] enforced more constraints on illumination and reflectance. In the work [4], Drew and Finlayson successfully improved the bilinear computation's efficiency by moving to logarithmic space.

The colour constancy problem mostly consists of estimating the illumination, or at least the chromaticity of the light in the observed scene. In [13], Robles-Kelly et al. generalized the assumptions and methods in colour constancy specifically for the IRSS problem. In principle, many colour constancy methods, such as [2, 9, 15], could be carried out in a multispectral domain, provided their underlying assumptions were still satisfied.

2.2 Conditional Random Field

Using a CRF model to incorporate the spatial information in a scene has been used in several works. Lu and Drew [10] applied a Markov Random Field (MRF) for shadow region segmentation. Their method first generated an illumination invariant image by using a model guided by the assumption of Planckian illumination. They formulated a pairwise potential which allowed them to label shadows smoothly. In [1], Beigpour and Riess et al. fused multiple physical and statistical based colour constancy methods into a CRF model which produced pixel-wise labels for shadows. However, their method relied on many complex colour constancy methods to generate accurate estimations for the CRF model, and this is time consuming process.

3. Low Rank Matrix Factorization

Here we briefly recapitulate the method [17], pointing out its innovative part that arrives at an excellent method. Firstly, for a pixel at spectral band i in a hyperspectral image, its value d_i is proportional to the product of the illumination l_i and surface reflectance r_i; that is,

$$d_i = l_i r_i,\ 1 \le i \le m, \tag{1}$$

Here, m denotes the number of spectral bands (which is 57 in all our experiments), and a vector accounting for camera gains has been omitted. As well, spectral sensitivity has been pre-corrected to be unity for all spectral bands $i = 1..m$.

For a hyperspectral image with n pixels under spatially uniform illumination, the intensity d_{ij} of the j-th pixel for the i-th spectral band is

$$d_{ij} = l_i r_{ij},\ 1 \le i \le m,\ 1 \le j \le n, \tag{2}$$

In matrix form this reads

$$\underbrace{\begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}}_{D_{m\times n}} = \underbrace{\begin{bmatrix} l_1 & & \\ & \ddots & \\ & & l_m \end{bmatrix}}_{L_{m\times m}} \underbrace{\begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{bmatrix}}_{R_{m\times n}} \tag{3}$$

The above system is underconstrained, since the observation matrix D has mn constraints but with $m(n+1)$ variables in the diagonal illumination matrix L and the reflectance matrix R taken together. In order to solve this equation, more constraints must be added to illumination and reflectance.

Reflectance spectra can easily be shown to usually lie in a low dimensional linear subspace. Therefore, eq. (3) could be written in the form

$$D_{m\times n} = L_{m\times m} R_{m\times n} = L_{m\times m} B_{m\times s} C_{s\times n}, \tag{4}$$

where B and C denote the spectral bases and coefficients respectively, and s is the subspace dimensionality. In addition to the above physically meaningful factorization, it could also be factorized using partial singular value decomposition (SVD) as follows

$$D_{m\times n} = U_{m\times s} S_{s\times s} V^T_{n\times s} = UQQ^{-1}SV^T = (UQ)(Q^{-1}SV^T), \tag{5}$$

Here, Q is an arbitrary $s \times s$ invertible matrix. Comparing eqs. 4 with 5, the IRSS problem comes down finding a proper matrix Q such that

$$UQ = LB \tag{6}$$

Recall that B is the reflectance basis matrix learned via principal component analysis of a spectra dataset [11, 12]. So based on eq. 6, the illumination spectrum L can be solved for by minimizing the following equation:

$$\min_{L,Q} ||UQ - LB||^2_F,\ \text{s.t. } l_i \ge \max_j d_{ij},\ l_z \le \max_{i,j} d_{ij} \tag{7}$$

in which $l_i \ge \max_j d_{ij}, 1 \le j \le n, 1 \le i \le m$ form constraints that each illumination factor should be greater than or equal to its corresponding observed signal, given eq. (2), since the reflectance r should be less than or equal to one. And $l_z \le \max_{i,j} d_{ij}$ where z is the row index of the maximum value $\max_{i,j} d_{ij}$: it restricts the illumination's absolute scale. Since eq. (7) is a simple convex quadratic program (QP), L and Q can be solved for by iteratively updating L and Q.

After obtaining L, the accuracy of separation can be further improved by minimizing the following cost functions on the original low rank model equation, eq. (4):

$$\begin{aligned} &\min_{L,C} ||D - LBC||^2_F, \\ &\text{s.t. } 0 \le (BC)_{ij} \le 1,\ l_i \ge \max_j d_{ij},\ l_z \le \max_{i,j} d_{ij} \end{aligned} \tag{8}$$

After optimizing this equation, the illumination L is well resolved and reflectance spectra values R are equal to BC.

4. Improved Local Separation and Illumination Patch Segmentation

4.1 Improved local separation

The previous work [17] can be understood as incorporating surface reflectance basis coherence into the separation problem; in other words, the low dimension property for surface reflectance helps to separate illumination and reflectance. In this paper, we are trying to implement the method [17] for individual patches in an image. We found in fact that separation accuracy for different patches varies a good deal. When the patch contains many colours, the separation of illumination and reflectance is reasonable good; but when colour diversity is small accuracy is diminished. This observation is consistent with the experiments in [17], in which separation accuracy increases with more colours.

After carefully reexamining the method [17], we understood that this problem is caused by eq. (5). In that equation, D is decomposed into U, S, V via SVD. So U is heavily dependent on D and could be noisy if D only contains small variation. For example we typically assume the reflectance subspace dimension s is 8 (chosen as the good tradeoff between expressive power and noise resistance), while on a small patch D could be well modelled only using only a 3-dimensional subspace; then U is noisy for the 5-D orthogonal subspace from SVD factorization.

Once the decomposed U is noisy, the next equations (7,8) would result in very poor separation. In order to resolve this issue of poor separation on local patches, we propose incorporating local spatial information into the final optimization equation (8). Our novel suggestion is to change (8) in the following way:

$$\begin{aligned} &\min_{\mathbf{l},C} ||D - LBC||^2_F + \lambda ||CG||^2_F, \\ &\text{s.t. } 0 \le (BC)_{ij} \le 1,\ l_i \ge \max_j d_{ij},\ l_z \le \max_{i,j} d_{ij} \end{aligned} \tag{9}$$

$$CG = \underbrace{\begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}}_{C_{m\times n}} \underbrace{\begin{bmatrix} & \vdots & \vdots & \vdots & \vdots & \\ \cdots & g_{n\_l} & g_{n\_u} & g_{n\_r} & g_{n\_b} & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}}_{G_{n\times 4n}}$$

Here G is a $n \times 4n$ sparse matrix; g is a n-by-1 column vector such as $[\cdots, 1, -1, \cdots]^T$; λ is an empirical parameter which we chose

as 0.05 in all our experiments. Thus CG represents the difference between the patch n and its neighbouring patch (left, up, right or bottom patch). Of course, here we considered the fact that border patches will not have all four neighbouring patches, and we set these differences to zero. By adding the latter loss term in equation (9), we are forcing the local patches to change smoothly. Hence we can alleviate the influence of a noisy U. In Table 1, we show a Peak Signal to Noise Ratio (PSNR) comparison between our method and method [17]. The PSNR is computed from the separated reflectance image and the ground truth original reflectance image.

Table 1. Improved local separation

Experiments	PSNR	RMS
Original	24.832	0.0573
Improved$^{\lambda=0.05}$	26.476	0.0474

4.2 Local illumination search

The main idea of our method is to divide images into patches, compute local estimations for each patch, and then unify patches into different illuminated regions using CRF. Then finally, we run again the local estimation method on regions to get a global estimation for different regions. This pipeline allows us to incorporate spatial information into the segmentation for different illuminated regions.

In order to test our algorithm, we synthesized test images by multiplying reflectance image with daylight illumination. The reflectance image data comes from dataset [8], which contains 22 multispectral reflectance images. We chose the bands from 420nm to 700nm and interpolated at an interval of 5nm.

For the illumination, we synthesized spectra using Judd's spectral power distribution (SPD) function $S(\lambda) = S_0(\lambda) + M_1 S_1(\lambda) + M_2 S_2(\lambda)$ to generate 56 daylight illuminations ranging from colour temperature 4000k to 25000k (4000:100:8500, 9000:500:10000, 11000:1000:15000, 20000, 25000). All daylight bands are the same for the multispectral reflectance image data, again ranging from 420nm to 700 with a interval of 5nm. There are in total 57 bands for each spectrum.

Crucially, as our method is aimed at estimating multiple illuminations in one image, we illuminated the image using two different illuminations, as depicted in Fig. (1b).

The first consideration for our method is to divide an image into patches. There are many options to segment images for this purpose. The most straightforward method is to uniformly divide each image into fixed-size patches. Another method is to use the superpixels segmentation method, which tends to segment at object boundaries and significant illumination changes. According to [5, 16, 6], these two methods result in almost the same accuracy much of the time. Hence, here we chose to use the uniform division method. One other reason for this choice is that the local estimation method [17] performs better when there are more colour variations. And uniformly dividing provides more diverse colour in each patch and we expect this to persist upon segmentation.

After dividing image into patches, we used the method described in sec. (4.1) to estimate illuminations for each patch. Here we add a small modification to this method. Instead of first factorizing the observed image using eq. (5), and then solving for illumination L using eq. (7), and finally optimizing illumination L and reflectance R via eq. (9). Here we simply loop through the possible illuminations from 4000k to 25000k, and find C for each illumination that will minimize equation (9). The illumination that has the minimum loss for eq. (9) is chosen as the estimated illumination for that patch. Fig. (1d) is the local estimation result for Fig. (1a), and we are showing the illuminations' indexes in Fig. (1d).

4.3 Different illumination regions segmentation via CRF

As can be seen from Fig. (1d), intuitively a pattern to segment between left and right appears apart from that based on the illumination. While if simply use a threshhold (e.g. 30 in our experiment) to classify the (1d) illumination map, we will obtain the result in (1e). Though (1d) looks like it contains two parts, the segmentation map is very coarse. The reason in this scene is that if we only use the threshhold to control segmentation then actually we only use the local information for this task. And as we already seen, the local estimation could be very noisy when a patch contains little colour information. Next, we consider using global spatial information for our task.

In addition to assigning each patch label (illumination 1, illumination 2) based on the estimated illumination, the labels should vary smoothly in the same illuminated region, while changes should appear at boundaries for different illuminated regions. The fact that a patch's label depends on the labels of its neighbouring patches allows us to model this optimizing problem as a CRF. The CRF based segmentation model is defined by meaningful spatial relationship between patches.

Here, the CRF model formulation is similar to other segmentation tasks. Here, we only consider the first order random field, which means only the neighbouring edges between patches are considered. Again, we denotes the multispectral image patches set as P, the edges set as N, and each patch takes a label l_p from $L = \{1,2\}$, with $\{1,2\}$ denoting illumination 1 and illumination 2.

Now we use the graph cut model in the Undirected Graphical Models (UGM) toolbox [14] to formulate our CRF problem, and solve for the best labels for each patch. In order to fit our model into UGM toolbox formulation, we wrote the equation in the following form:

$$Potential(l) = \sum_{p \in P} U_p(l_p) + \sum_{(p,q) \in N} B(l_p, l_q), \tag{10}$$

where U_p denotes the unary potentials to assign patch p to illumination 1 or illumination 2; larger values of U denote higher confidence that p belongs to that illumination. And $B(l_p, l_q)$ denotes the binary potentials for assigning patch p and q to individual illumination (they could be the same or not). This binary potential represents the confidence with which we believe p and q should be assigned the same label or not.

In our method, we did not see a clear cue to indicate the likelihood for assigning each patch to illumination 1 or 2. So we assign equal potentials to $U_p = \exp^{w_{constant}}$ ($w_{constant} = 0.015$ in our experiment), and therefore we actually do not use the first term in the equation (10).

Now the task is to find the graph cut model for illumination 1 and illumination 2 that maximizes the potential in eq. (10). The

binary potential function $B(l_p, l_q)$ is defined as:

$$B(l_p, l_q) = \begin{cases} 1 & \text{if } l_p \neq l_q \\ \exp^{w_1 + w_2 * (|Idx_p - Idx_q|/K)} & \text{if } l_p \equiv l_q \end{cases} \quad (11)$$

where $K = \max_{\{p,q\} \in N} (|Idx_p - Idx_q|)$ and Idx_p, Idx_q are the estimated illumination indexes from method sec. (4.1) . Intuitively, when p and q are illuminated in the same region, B should be larger when assigning $l_p \equiv l_q$ than for $l_p \neq l_q$. In general, if neighbouring patches' local estimation results are similar, then they would have high probability to be illuminated in the same region, and thus we should assign high potentials to $B(l_p, l_q)$ for labeling them the same, and vice versa.

Throughout our experiments, we chose $w_1 = 2.8$ and $w_2 = -1.9$. And this model satisfied our analysis above very well, since when Idx_p and Idx_q are similar, then the potential for B is larger when their labels are the same. Also when Idx_p and Idx_q are different, B is larger when their labels are different.

Finally finding the best labels is finding maximum potentials for eq. (10) in this model. Here, the reader is referred the the toolbox [14] for more details about UGM modelling for the graph-cut algorithm. Using the graph-cut decoding algorithm from [14], we segmented the image into different illuminated regions, as shown in Fig. (1f).

5. Illumination and Reflectance Separation

Once the illumination region is segmented, we again run our method described in sec. (4.2) on each region to find the best illumination candidate. As we have then solved for the illumination for each region, the reflectance can this be solved for by using the observed image dividing the illumination. In Fig. (2), we show the separation result of [17] and of our method.

As can be seen from Fig. (2d), method [17] almost get the illumination totally wrong. We firstly expect the result would be the average of two illuminations, while as the result is only trying minimizing equation 7, there is no cue for how to balance between these two illuminations. And from Fig. (2b), we can see the recovered reflectance differs from the original reflectance image, as well as there are two obvious region in the image (which supposed to be uniform like the original reflectance image).

For our method, since we assumed there are two illuminations in the scene, then we can see the recovered two illuminations in Fig. (2e) fits the true illumination quite well. And more surprisingly, the recovered reflectance image in Fig. (2c) looks quite uniform and close to the original reflectance image. Of course, we note that the wrongly segmented patches in Fig. (1f) are the noisy part in our final recovered reflectance image.

6. Conclusion

From the experiments given, we have demonstrated that dividing the image into patches helps for multi-illumination estimation, and using a CRF model greatly alleviates the noise introduced by local estimation.

Note that our model is not restricted to two illuminations conditions. In eq. (10), the labels state could be larger than 2. By simply changing the potential functions eq. (11) for more states, our model could handle more illumination conditions. However, of course this extension would definitely introduce more noise into the segmentations.

Acknowledgement

One of the authors (X.C.) thanks Prof. Imari Sato and Dr. Yinqiang Zheng for helpful discussions during his internship in NII, Tokyo, Japan.

References

[1] Shida Beigpour, Christian Riess, Joost Van De Weijer, and Elli Angelopoulou. Multi-illuminant estimation with conditional random fields. *Image Processing, IEEE Transactions on*, 23(1):83–96, 2014.

[2] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980.

[3] Po-Rong Chang and Tsung-Hsieh Hsieh. Constrained nonlinear optimization approaches to color-signal separation. *IEEE Transactions on Image Processing*, 4(1):81–94, Jan 1995.

[4] Mark S. Drew and Graham D. Finlayson. Analytic solution for separating spectra into illumination and surface reflectance components. *J. Opt. Soc. Am. A*, 24(2):294–303, Feb 2007.

[5] Pedro F Felzenszwalb and Daniel P Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.

[6] A. Gijsenij, R. Lu, and T. Gevers. Color constancy for multiple light sources. *IEEE Transactions on Image Processing*, 21(2):697–707, Feb 2012.

[7] J. Ho, B. V. Funt, and M. S. Drew. Separating a color signal into illumination and surface reflectance components: theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(10):966–977, Oct 1990.

[8] S. Hordley, G. Finalyson, and P. Morovic. A multi-spectral image database and its application to image rendering across illumination. In *Image and Graphics (ICIG'04), Third International Conference on*, pages 394–397, Dec 2004.

[9] Edwin H Land and John J McCann. Lightness and retinex theory. *JOSA*, 61(1):1–11, 1971.

[10] Cheng Lu and Mark S Drew. Shadow segmentation and shadow-free chromaticity via markov random fields. In *Color and Imaging Conference*, volume 2005, pages 125–129. Society for Imaging Science and Technology, 2005.

[11] Laurence T Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA A*, 3(10):1673–1683, 1986.

[12] Jussi PS Parkkinen, J Hallikainen, and T Jaaskelainen. Characteristic spectra of munsell colors. *JOSA A*, 6(2):318–322, 1989.

[13] Antonio Robles-Kelly and Cong Phuoc Huynh. *Imaging spectroscopy for scene analysis*. Springer Science & Business Media, 2012.

[14] M. Schmidt. Ugm: A matlab toolbox for probabilistic undirected graphical models, 2007.

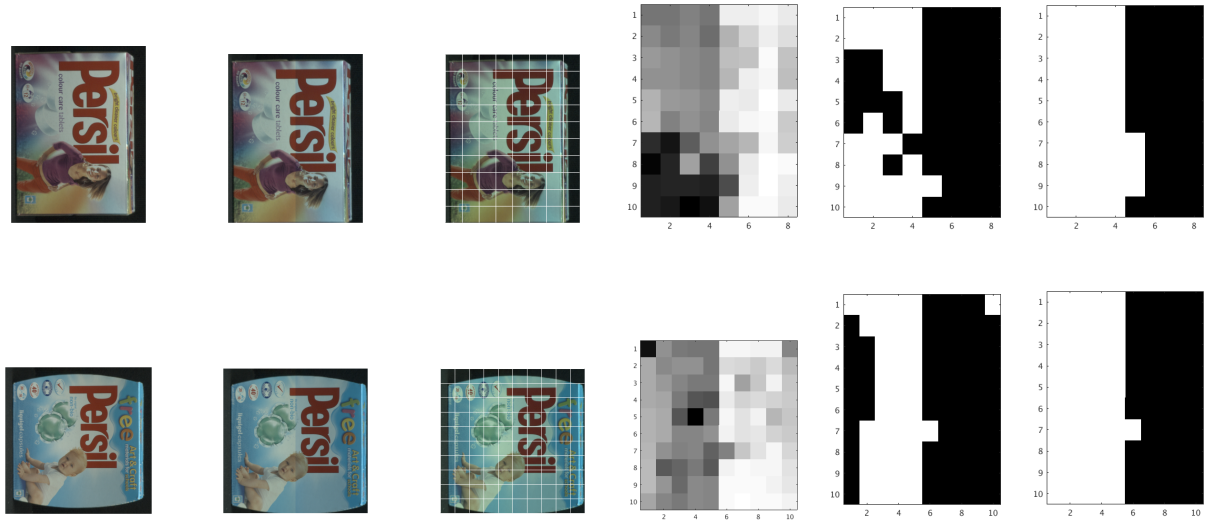
[15] Robby T Tan, Ko Nishino, and Katsushi Ikeuchi. Illumination chromaticity estimation using inverse-intensity chromaticity space. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 1, pages I–673. IEEE, 2003.

[16] Olga Veksler, Yuri Boykov, and Paria Mehrani. Superpixels and supervoxels in an energy optimization framework. In *Computer Vision–ECCV 2010*, pages 211–224. Springer, 2010.

[17] Y. Zheng, I. Sato, and Y. Sato. Illumination and reflectance spectra separation of a hyperspectral image meets low-rank matrix factorization. In *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, pages 1779–1787, June 2015.

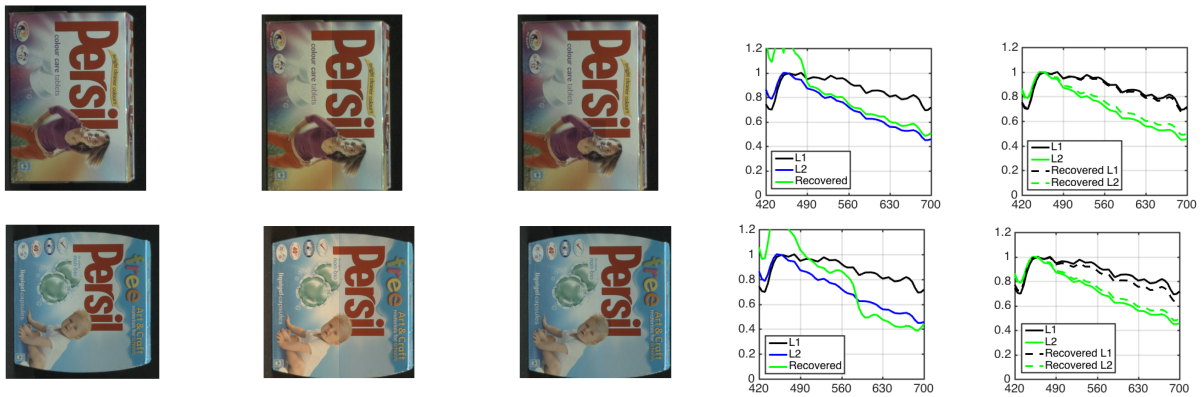
Author Biography

Mark S. Drew is a Professor in the School of Computing Science at Simon Fraser University in Vancouver, Canada. His background education is in Engineering Science, Mathematics, and Physics. His interests lie in the fields of image processing, color, computer vision, computer graphics, multimedia, and visualization. He has published over 160 refereed papers and is the holder of 9 patents and applications in computer vision, image processing, and image reconstruction and accenting.



(a) Reflectance Image (b) 2 illuminations (c) Uniformly divided (d) Local estimation (e) Simple classify (f) Graph cut result

Figure 1: All the displaying image are hyperspectral images, we convert them to RGB image for better display. (a) is the original reflectance image, (b) is the reflectance image illuminated by 4000K daylight on the left half and 8000K daylight on the right half, (c) is the way we divide image into patches, each patch size is 32-by-32, (d) is showing the illumination's index ranging from 1 to 56, (e) is the binary classification result when use index 30 as threshhold to segment the image, (f) is the segmented different illuminated regions using graph cut based on local estimation



(a) Reflectance Image (b) [17] recovered (c) Ours recovered (d) [17] recovered (e) Ours recovered

Figure 2: (a) is the original reflectance image, (b) is the recovered reflectance image based on method [17] which assume one illumination in the scene, (c) is the recovered reflectance using our method, (d) is recovered illumination using method [17], (e) is recovered illumination using our method