Image Segmentation with Simultaneous Illumination and Reflectance Estimation: An Energy Minimization Approach

Chunming Li[1], Fang Li[2], Chiu-Yen Kao[3], Chenyang Xu[4]

[1] Vanderbilt University Institute of Imaging Science, Nashville, TN 37232, USA
[2] Dept. of Mathematics, East China Normal University, Shanghai 200062, China
[3] Dept. of Mathematics, The Ohio State University, Columbus, OH 43210, USA
[4] Siemens Corporate Research, Princeton, NJ 08540, USA
lchunming@gmail.com

Abstract

Spatial intensity variations caused by illumination changes have been a challenge for image segmentation and many other computer vision tasks. This paper presents a novel method for image segmentation with simultaneous estimation of illumination and reflectance images. The proposed method is based on the composition of an observed scene image with an illumination component and a reflectance component, known as intrinsic images. We define an energy functional in terms of an illumination image, the membership functions of the regions, and the corresponding reflectance constants of the regions in the scene. This energy is convex in each of its variables. By minimizing the energy, image segmentation result is obtained in the form of the membership functions of the regions. The illumination and reflectance components of the observed image are estimated simultaneously as the result of energy minimization. With illumination taken into account, the proposed method is able to segment images with non-uniform intensities caused by spatial variations in illumination. Comparisons with the state-of-the-art piecewise smooth model demonstrate the superior performance of our method.

1. Introduction

An observed image of a scene is a composition of illumination and reflectance components, which have been termed as *intrinsic images* in [1]. The reflectance component characterizes a unique physical property of the object surfaces in the scene – the *albedo* of the object surfaces. The illumination component depends on the light direction and the orientation of the object surfaces, and can be used to infer the geometry of the objects being viewed. However, spatial variations in illumination cause changes of the intensities in the observed images, which has been a challenge for many computer vision tasks, such as segmentation, tracking, and object recognition.

In image segmentation, a considerable difficulty arises from spatial variations in illumination. The observed image intensities in a single object may not be uniform due to spatial variations in illumination. In this case, it is difficult to define a region descriptor in region-based methods for image segmentation. Therefore, many widely used region-based methods [15, 5] rely on the uniformity of the intensities in the regions to be segmented, as it is easy to model the intensities within the regions of interest and to define a region descriptor. These methods are not applicable to real-world images with non-uniform intensities caused by spatial variations in illumination.

The well-known Mumford-Shah model in [11] has provided a theoretical framework of image segmentation, which can be applied to a wider range of images, including those with non-uniform intensities. The Mumford-Shah model assumes that an image can be approximated by a piecewise smooth function. Based on this assumption, the Mumford-Shah model seeks a contour and a number of smooth functions that approximate the image intensities in disjoint regions separated by the contour. The smoothness of these functions is ensured by imposing a smoothing term in the Mumford-Shah model, which leads to a set of partial differential equations that have to be solved at each iteration for the evolution of the contour. Therefore, the computation in these methods is rather expensive, which limits their practical applications. In addition, due to the non-convexity of the underlying energy functionals, the corresponding energy minimization algorithms may converge to false local minima.

The piecewise smooth characterization of images in the Mumford-Shah model is too general for real images. A

2009 IEEE 12th International Conference on Computer Vision (ICCV)
978-1-4244-4419-9/09/$25.00 ©2009 IEEE

more specific characterization of real images would provide more useful information, which can be taken into account in image segmentation. In fact, there is a more specific characterization of real images based on the theory of lightness perception: an observed image I, its illumination image component S, and reflectance image component R are related by the following multiplicative model:

$$I = R \cdot S. \tag{1}$$

Furthermore, the illumination image S is assumed to be smooth [6, 8, 2]. The reflectance image R can be approximated by a piecewise constant map, as it describes a physical property of the object surfaces in the scene. Therefore, the observed image can be approximated by the product of a smooth function and a piecewise constant function.

In this paper, we propose a novel energy minimization method for image segmentation with simultaneous estimation of illumination and reflectance images, based on the above image composition model in Eq. (1) and the properties of the illumination and reflectance images. We define an energy functional in terms of an illumination image, the membership functions of the regions, and the corresponding reflectance constants of the regions in the scene. By minimizing the energy, image segmentation result is obtained in the form of the membership functions of the regions. Meanwhile, the illumination and reflectance components of the observed image are estimated simultaneously as the result of energy minimization.

The proposed method is related to the techniques for *illumination and reflectance estimation (IRE)* from an observed image, such as those in [7, 14, 9, 12]. However, our method is fundamentally different from these methods for illumination and reflectance estimation. Our method is mainly targeted for image segmentation. In the formulation of our method, we introduce membership functions of the object surfaces in the image, which directly represent an image segmentation result. The existing methods for IRE do not provide image segmentation results.

2. Problem Formulation Based on Intrinsic Image Composition

For theoretical completeness, we slightly extend the image model (1) with additive noise as below:

$$I = R \cdot S + n \tag{2}$$

where n is assumed to be zero-mean Gaussian noise. The reflectance image R is approximately a constant r_i in a region Ω_i, which corresponds to an object surface in the scene. The property of the illumination image S and the reflectance image R mentioned in Section 1 can be made more specific as below:

(A1) $R(\mathbf{x}) \approx r_i$ for $\mathbf{x} \in \Omega_i$ for $i = 1, \cdots, N$, with $\Omega_i \cap \Omega_j = \emptyset$ for any $i \neq j$ and $\cup_{i=1}^N \Omega_i = \Omega$.

(A2) The illumination image S is smooth.

The above disjoint regions $\Omega_1, \cdots, \Omega_N$ form a partition of the image domain Ω. The goal of image segmentation is to determine such a partition.

From the above image model (2), we propose to seek a piecewise constant map $\hat{R}$ and a smooth function $\hat{S}$ to estimate the intrinsic images R and S, respectively. In general, a piecewise constant map R can be expressed as a linear combination:

$$R = \sum_{i=1}^{N} r_i \chi_{\Omega_i} \tag{3}$$

where χ_{Ω_i} is the characteristic function of the region Ω_i. Each characteristic function χ_{Ω_i} is a membership function of the region Ω_i. Thus, estimation of the reflectance image R from an observed image I can be achieved by seeking membership functions u_i associated with the regions Ω_i and the corresponding reflectance constants r_i, for $i = 1, \cdots, N$. Thus, image segmentation result is given by the obtained membership functions.

For convenience, the membership functions $u_1, \cdots, u_N$ and the constants $r_1, \cdots, r_N$ are denoted by a vector valued membership function $U = (u_1, \cdots, u_N)$, and a vector $\mathbf{r} = (r_1, \cdots, r_N)$, respectively. Thus, the problem of estimating the intrinsic images R and S can be formulated as one that seeks optimal vector membership function U, vector $\mathbf{r}$, and a smooth function S according to an optimality criterion. We will propose a variational formulation, in which the optimal U, $\mathbf{r}$, and S are obtained by minimizing an energy functional $\mathcal{F}(U, \mathbf{r}, S)$.

3. Energy Formulation

3.1. Statistical Property of Local Intensities

The variation in appearance caused by the illumination changes has been a challenging problem for visual tasks, such as segmentation. However, the smoothness of illumination image, along with the piecewise constant property of the reflectance image, implies a useful property of the observed image intensities in local area, which will be used in the formulation of the proposed energy minimization framework for segmentation with simultaneous illumination and reflectance estimation. In fact, the observed image intensities in a small local area follow a simple distribution as described below.

We consider a circular neighborhood of each point $\mathbf{x} \in \Omega$, defined by $\mathcal{O}_{\mathbf{x}} \triangleq \{\mathbf{y} : |\mathbf{y} - \mathbf{x}| \leq \rho\}$, where ρ is the radius of the neighborhood. The partition $\{\Omega_i\}_{i=1}^N$ of the entire domain Ω induces a partition of the neighborhood $\mathcal{O}_{\mathbf{x}}$ into N subsets $\mathcal{O}_{\mathbf{x}} \cap \Omega_i$, which form a partition of $\mathcal{O}_{\mathbf{x}}$. For a smooth function S, the values $S(\mathbf{y})$ for all $\mathbf{y}$ in the circular neighborhood $\mathcal{O}_\mathbf{x}$ can be well approximated by $S(\mathbf{x})$. Therefore, the intensities $S(\mathbf{y})R(\mathbf{y})$ in each subregion $\mathcal{O}_\mathbf{x} \cap \Omega_i$ are approximately the constant $S(\mathbf{x})r_i$. Thus, we have the following approximation

$$S(\mathbf{y})R(\mathbf{y}) \approx S(\mathbf{x})r_i \quad \text{for } \mathbf{y} \in \mathcal{O}_\mathbf{x} \cap \Omega_i \tag{4}$$

From the image model (2), we have

$$I(\mathbf{y}) \approx S(\mathbf{x})r_i + n(\mathbf{y}) \quad \text{for } \mathbf{y} \in \mathcal{O}_\mathbf{x} \cap \Omega_i$$

where $n(\mathbf{y})$ is additive zero-mean Gaussian noise.

The above arguments show that the intensities within the neighborhood $\mathcal{O}_\mathbf{x}$ are samples from N Gaussian distributions with distinct means $S(\mathbf{x})r_i$, $i = 1, \cdots, N$. In the other word, the intensities in the neighborhood $\mathcal{O}_\mathbf{x}$ form N clusters:

$$\{I(\mathbf{y}) : \mathbf{y} \in \mathcal{O}_\mathbf{x} \cap \Omega_i\},$$

each with a cluster center $m_i \approx S(\mathbf{x})r_i$. As shown above, the cluster center m_i can be approximated by the product of a region dependent reflectance constants r_i and an illumination factor $S(\mathbf{x})$, which approximates the illumination within the neighborhood $\mathcal{O}_\mathbf{x}$. This property of local intensities is fully exploited in the proposed method for image segmentation and estimation of intrinsic images.

3.2. Clustering Criterion for Local Intensities

In view of the above clustering characterization of local intensities in terms of illumination and reflectance, we introduce a clustering-based approach for image segmentation with simultaneous illumination and reflectance estimation. Lets first review a basic clustering criterion that has been used to derive the well known fuzzy C-means and K-means algorithms. A standard clustering algorithm seeks the membership functions u_i and the cluster centers c_i, such that the following clustering criterion function is minimized:

$$J_{\text{FCM}} = \int \sum_{i=1}^{N} u_i^q(\mathbf{x}) |I(\mathbf{x}) - c_i|^2 d\mathbf{x} \tag{5}$$

where $q \geq 1$ is the fuzzifier, I is the input image, $u_i(\mathbf{x})$ is the membership value at point $\mathbf{x}$ for class i such that $\sum_i^N u_i(\mathbf{x}) = 1$, and c_i is the cluster center of class i. This clustering criterion has been used to derive a simple yet powerful algorithm to classify data that consists of N separable clusters with distinct cluster centers and determine the cluster centers at the same time. The corresponding minimization process is the well-known fuzzy C-means (FCM) algorithm. As a special case when the fuzzifier q is 1, the minimization of the above clustering criterion is achieved by the K-means algorithm. The above FCM clustering criterion is in general not valid for image I with a spatially varying illumination image component S. However, from the statistics of local intensities in Section 3.1, the above standard clustering criterion can be applied for local intensity classification.

The analysis in Section 3.1 has shown that the intensities in the neighborhood $\mathcal{O}_\mathbf{x}$ form N clusters, each with a distinct cluster center $m_i \approx S(\mathbf{x})r_i$. Therefore, we propose to minimize the following clustering criterion for local intensities in the neighborhood $\mathcal{O}_\mathbf{x}$:

$$\mathcal{J}_\mathbf{x}^{loc}(U, \mathbf{r}, S(\mathbf{x})) \triangleq \sum_{i=1}^{N} \int_{\mathcal{O}_\mathbf{x}} u_i^q(\mathbf{y}) K(\mathbf{x}-\mathbf{y}) \cdot |I(\mathbf{y}) - S(\mathbf{x})r_i|^2 d\mathbf{y} \tag{6}$$

where u_i is the membership function of the region Ω_i, and $K(\mathbf{x}-\mathbf{y})$ is a weighting function, which decreases as the distance from $\mathbf{y}$ to the neighborhood center $\mathbf{x}$ increases, and becomes zero when $|\mathbf{x}-\mathbf{y}| > \rho$, i.e., $\mathbf{y} \notin \mathcal{O}_\mathbf{x}$. In this paper, the notation $|a|$ represents the Euclid norm for a vector a or the absolute value for a scalar a. The above criterion function can be used for both gray level images and color images.

The above clustering criterion function $\mathcal{J}_\mathbf{x}^{loc}$ can be written as

$$\mathcal{J}_\mathbf{x}^{loc}(U, \mathbf{r}, S(\mathbf{x})) = \sum_{i=1}^{N} \int u_i^q(\mathbf{y}) K(\mathbf{x}-\mathbf{y}) \cdot |I(\mathbf{y}) - S(\mathbf{x})r_i|^2 d\mathbf{y} \tag{7}$$

due to the fact that $K(\mathbf{x}-\mathbf{y}) = 0$ for $\mathbf{y} \notin \mathcal{O}_\mathbf{x}$.

3.3. Integration of Local Clustering Criterion

The optimal U, $\mathbf{r}$, and S are defined as those that minimize the clustering criterion function $\mathcal{J}_\mathbf{x}^{loc}(U, \mathbf{r}, S(\mathbf{x}))$ for all the neighborhood $\mathcal{O}_\mathbf{x}$. Minimization of $\mathcal{J}_\mathbf{x}^{loc}$ for all $\mathbf{x} \in \Omega$ can be achieved by minimizing the integral of $\mathcal{J}_\mathbf{x}^{loc}$ over Ω. Therefore, we define an energy $\mathcal{J}(\mathbf{u}, \mathbf{r}, S) \triangleq \int \mathcal{J}_\mathbf{x}^{loc}(\mathbf{u}, \mathbf{r}, S(\mathbf{x})) d\mathbf{x}$, i.e.

$$\mathcal{J}(U, \mathbf{r}, S) \triangleq \int \sum_{i=1}^{N} \int u_i^q(\mathbf{y}) K(\mathbf{x}-\mathbf{y}) \cdot |I(\mathbf{y}) - S(\mathbf{x})r_i|^2 d\mathbf{y} d\mathbf{x} \tag{8}$$

By changing the order of integration, the above energy $\mathcal{J}(\mathbf{u}, \mathbf{r}, S)$ can be written in the form:

$$\mathcal{J}(U, \mathbf{r}, S) = \int \sum_{i=1}^{N} u_i^q(\mathbf{y}) d_i(I(\mathbf{y})) d\mathbf{y} \tag{9}$$

where

$$d_i(I(\mathbf{y})) \triangleq \int K(\mathbf{x}-\mathbf{y}) |I(\mathbf{y}) - S(\mathbf{x})r_i|^2 d\mathbf{x} \tag{10}$$

Image segmentation and the estimation of illumination and reflectance can be achieved by solving the following constrained energy minimization problem:

$$\min \mathcal{J}(U, \mathbf{r}, S) \quad \text{subject to } U \in \mathcal{U}, \tag{11}$$

where $\mathcal{U}$ is the space of all the membership functions, i.e.

$$\mathcal{U} \triangleq \{(u_1, \cdots, u_N)^T : 0 \leq u_i(x) \leq 1, i = 1, \cdots, N, \text{ and } \textstyle\sum_{i=1}^N u_i(x) = 1, \text{ for all } x \in \Omega\} \tag{12}$$

3.4. Energy Formulation with Membership Regularization

For images with high level noise, it is necessary to regularize membership function by adding a regularization term $\mathcal{R}(U)$ to the above clustering energy $\mathcal{J}(U, \mathbf{r}, S)$. Therefore, we define

$$\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S) = \mathcal{J}(U, \mathbf{r}, S) + \gamma \mathcal{R}(U) \tag{13}$$

In this work, we use total variation of the membership functions as the *membership regularization term*, i.e.

$$\mathcal{R}(U) = \sum_{i=1}^{N} \int |\nabla u_i(\mathbf{x})| d\mathbf{x} \tag{14}$$

Then we define

$$\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S) \triangleq \mathcal{J}(U, \mathbf{r}, S) + \gamma \mathcal{R}(U), \tag{15}$$

and solve the following constrained energy minimization problem:

$$\min \mathcal{J}^{\text{reg}}(U, \mathbf{r}, S) \quad \text{subject to } U \in \mathcal{U}, \tag{16}$$

It is worth noting that this energy is convex in the variables $\mathbf{r}$ and S, which implies a unique global minimum in each of these variables given the others. The minimization of the energy functionals $\mathcal{J}(U, \mathbf{r}, S)$ and $\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S)$ is described in the next section.

4. Energy Minimization

We have defined two constrained energy minimization problems in Eqs. (11) and (16) in the previous section. The energy minimization for the variables $\mathbf{r}$ and S are the same for both cases, as the additional membership regularization term $\gamma\mathcal{R}(U)$ in (16) is fixed as a constant when minimizing the energy functionals $\mathcal{J}(U, \mathbf{r}, S)$ and $\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S)$with respect to $\mathbf{r}$ and S.

4.1. Minimization of $\mathcal{J}(U, \mathbf{r}, S)$

4.1.1 Case $q > 1$

Minimization of the energy $\mathcal{J}(U, \mathbf{r}, S)$ can be performed by interleaving the minimization with respect to the variables U, $\mathbf{r}$, and S. Note that the energy $\mathcal{J}(U, \mathbf{r}, S)$ is convex in each variable, and it is a quadratic form in terms of the variables $\mathbf{r}$ and S. The minimization with respect to each variable, given the other two fixed, can be achieved by a closed form solution as below:

- For fixed U and S, there is a unique minimizer of the energy $\mathcal{J}(U, \mathbf{r}, S)$ with respect to $\mathbf{r}$, which is in the form denoted by $\hat{\mathbf{r}} = (\hat{r}_1, \cdots, \hat{r}_N)$, is given by

$$\hat{r}_i = \frac{\int (S * K) I u_i^q d\mathbf{x}}{\int (S^2 * K) u_i^q d\mathbf{x}}, \quad i = 1, \cdots, N. \tag{17}$$

where $*$ is the convolution operation.

- For fixed $\mathbf{r}$ and S, there is a unique minimizer of the energy $\mathcal{J}(U, \mathbf{r}, S)$ with respect to U, which is given by

$$\hat{u}_i(\mathbf{y}) = \frac{1}{\sum_{k=1}^N \left(\frac{d_i(I(\mathbf{y}))}{d_k(I(\mathbf{y}))}\right)^{\frac{1}{q-1}}} \tag{18}$$

- Given U and $\mathbf{r}$, there is a unique minimizer of the energy $\mathcal{J}(U, \mathbf{r}, S)$ with respect to S, which is given by

$$\hat{S} = \frac{(IR^{(1)}) * K}{R^{(2)} * K} \tag{19}$$

where $R^{(1)} = \sum_{i=1}^N r_i u_i^q$ and $R^{(2)} = \sum_{i=1}^N r_i^2 u_i^q$.

Note that the convolutions with a kernel K in the expression of $\hat{S}$ in Eq. (19) confirms the smoothness of the derived optimal illumination image $\hat{S}$, which has been explained from the definition of the clustering criterion function $\mathcal{J}_{\mathbf{x}}^{loc}$ in the previous section.

4.1.2 Case $q = 1$

For the case of $q = 1$, the optimization of the variables $\mathbf{r}$ and S in the energy $\mathcal{J}(U, \mathbf{r}, S)$ is the same as (17) and (19) with $q = 1$. However, the optimal membership function U cannot be computed by (18) in this case. The optimization of $\mathbf{r}$ for the case of $q = 1$ is given below.

Let $u_1, \cdots, u_N$ be positive functions such that $\sum_{i=1}^N u_i(\mathbf{x}) = 1$ for all $\mathbf{x} \in \Omega$, then the energy

$$\int \sum_{i=1}^{N} d_i(I(\mathbf{x})) u_i(\mathbf{x}) d\mathbf{x} \tag{20}$$

is minimized if for each $\mathbf{x} \in \Omega$, the values of $u_1(\mathbf{x}), \cdots, u_N(\mathbf{x})$ are given by

$$\hat{u}_i(\mathbf{x}) = \begin{cases} 1, & i = i_{\min}(\mathbf{x}); \\ 0, & i \neq i_{\min}(\mathbf{x}). \end{cases} \tag{21}$$

where

$$i_{\min}(\mathbf{x}) = \arg\min_i \{d_i(I(\mathbf{x}))\}.$$

4.2. Minimization of $\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S)$ with Respect to U

Using the same technique as in [10]. The first constraint can be relaxed by adding exact penalty convex functions $\nu(u_i) = \max(0, |2u_i - 1| - 1)$. Using Lagrange multiplier method, the second constraint can be removed by adding a term $\alpha \int (\sum_{i=1}^N u_i - 1)^2 d\mathbf{x}$ in (15) where α is a positive Lagrange multiplier. Then the equivalent unconstrained problem is to minimize

$$\mathcal{J}^{\text{reg}}(U, \mathbf{r}, S) = \sum_{i=1}^N \left(\gamma \int |\nabla u_i| d\mathbf{x} + \int u_i d_i d\mathbf{x} + \nu(u_i)\right) + \frac{\alpha}{2} \int (\sum_{i=1}^N u_i - 1)^2 d\mathbf{x}. \quad (22)$$

Observing that the energy (22) is convex in variables u_i, we choose to follow [3] and take use of the fast duality projection method in [4]. So we add N auxiliary variables v_i and minimize the following approximate energy

$$\mathcal{J}^{\text{reg}}(U, V, \mathbf{r}, S) = \sum_{i=1}^N (\gamma \int |\nabla v_i| d\mathbf{x} + \frac{1}{2\theta} \int (v_i - u_i)^2 d\mathbf{x} + \int u_i d_i d\mathbf{x} + \nu(u_i)) + \frac{\alpha}{2} \int (\sum_{i=1}^N u_i - 1)^2 d\mathbf{x} \quad (23)$$

where $\theta > 0$ is chosen to be small enough so that u_i and v_i are almost identical with respect to the L^2 norm. Since (23) is convex in u_i and v_i, we can use alternate minimization method to find the minimizer of (23). Decompose the problem into four sub-problems as follows

- For fixed $U, \mathbf{r}, S$, solve v_i by minimizing

  $$\gamma \int |\nabla v_i| d\mathbf{x} + \frac{1}{2\theta} \int (v_i - u_i)^2 d\mathbf{x}$$

  This problem can be efficiently solved by fast duality projection algorithm. The solution is given by

  $$v_i = u_i - \theta \operatorname{div} p_i, i = 1, ..., K$$

  where the vector p_i can be solved by fixed point method: Initializing $p_i^0 = 0$ and iterating

  $$p_i^{n+1} = \frac{p_i^n + \tau \nabla (\operatorname{div} p_i^n - u_i/\theta)}{1 + \tau |\nabla (\operatorname{div} p_i^n - u_i/\theta)|}.$$

  with $\tau \leq 1/8$ to ensure convergence.

- For fixed $V, \mathbf{r}, S$, we solve u_i by minimizing

  $$\frac{1}{2\theta} \int (v_i - u_i)^2 d\mathbf{x} + \int u_i d_i d\mathbf{x} + \nu(u_i) + \frac{\alpha}{2} \int (\sum_{i=1}^N u_i - 1)^2 d\mathbf{x}.$$

  Remark that the presence of the $\nu(u_i)$ term in the energy is equivalent to cutting off each u_i at 0 and at 1. On the other hand, if $u_i \in [0, 1]$, then $\nu(u_i) = 0$, and the Euler-Lagrange equation of this problem is

  $$\frac{1}{\theta}(u_i - v_i) + d_i + \alpha(\sum_{i=1}^N u_i - 1) = 0$$

  Then the minimization is given by

  $$u_i = \min(\max(\frac{v_i - \theta d_i - \theta\alpha(\sum_{j \neq i} u_j - 1)}{1 + \theta\alpha}, 0), 1).$$

- For fixed U and V, $\mathbf{r}$ and S are given by Eqs. (17) and (19) respectively.

5. Implementation and Experimental Results

5.1. Implementation and Parameter Setting

The energy minimization is performed by an interleaved optimization of each variable of the energy $\mathcal{J}(U, \mathbf{r}, S)$ in an iterative process. In every iteration, we minimize the energy with respect to one of the three variables given the other two obtained from the previous iteration. Before the iteration, the three variables U, $\mathbf{r}$, and S are initialized as follows. In our implementation, we always initialize the illumination image S as a constant map for simplicity. To reduce the iteration number and save computation time, we apply the efficient K-means algorithm as a preliminary segmentation for the initialization of the membership function U and the reflectance constant $\mathbf{r}$.

After the optimal $\hat{U}$, $\hat{\mathbf{r}}$, and $\hat{S}$ are obtained as the result of energy minimization, the segmentation result is given by the membership functions in $\hat{U} = (\hat{u}_1, \cdots, \hat{u}_N)$. The illumination and reflectance images are given by $\hat{S}$ and

$$\hat{R}(x) = \sum_{i=1}^N \hat{r}_i \hat{u}_i(x) \quad (24)$$

To estimate the illumination image more accurately, it is preferable to choose a small scale parameter σ in the Gaussian kernel K. In this case, we use $\sigma = 1$ for the estimation of illumination and reflectance images. A larger scale parameter σ would blur the resulting illumination image. If one is only interested in the segmentation result, a large scale parameter σ can be used in our method. In addition, the fuzzifier q is set to $q = 1.2$ in most of our experiments.

5.2. Experimental Results

We first demonstrate the effectiveness of the proposed method for three images in Column 1 of Fig. 1. The spatial illumination variation is obvious in these images. The segmentation results and illumination images are shown in Columns 2 and 3, respectively. The segmentation results are highly consistent with the original images. The estimated reflectance images exhibit desirable homogeneity in each region, while intensity variation due to the illumination changes is extracted to the estimated illumination images in Column 3, which characterize the shapes of the surfaces in the scenes and the lighting condition.

Figure 1. Results for three color images. Column 1: Original image; Column 2: Segmentation result; Column 3: Estimated illumination image.

We have also applied our method to face images, as shown in Fig. 2. The variation in appearance caused by spatial illumination changes often poses additional challenge in face recognition. Therefore it is in general desirable to obtain an illumination free image, i.e., a pure reflectance image, as the input to face recognition systems to improve robustness and recognition rate. Fig. 2 shows the results of our method for two face images in Column 1. The segmentation results are shown in Column 2. Note that the skins on both faces and necks are correctly identified as a single region. As a result, the corresponding estimation of the reflectance images exhibit such homogeneity in the faces and necks. The spatial variation in illumination, characterizing the geometry of the face, is extracted in the illumination images, as shown in Column 3.

The proposed method is by nature able to deal with intensity inhomogeneities in image segmentation. Fig. 3 shows the segmentation results of our method for four images. There are fast local intensity variations, such as the intensity variations within the grass or the squirrel in the first image. For such images, it would be helpful to smooth the image to significantly suppress fast local intensity variations. Our method is then applied to the smoothed images, which gives desirable segmentation results as shown in Fig. 3.

As mentioned in Section 1, the Mumford-Shah model is able to deal with intensity in image segmentation. Naturally, we compare our method with the piecewise smooth (PS) model proposed in [13], which is a well-known method that

Figure 2. Result for two face images. Column 1: Original image; Column 2: Segmentation result; Column 3: Estimated illumination image.

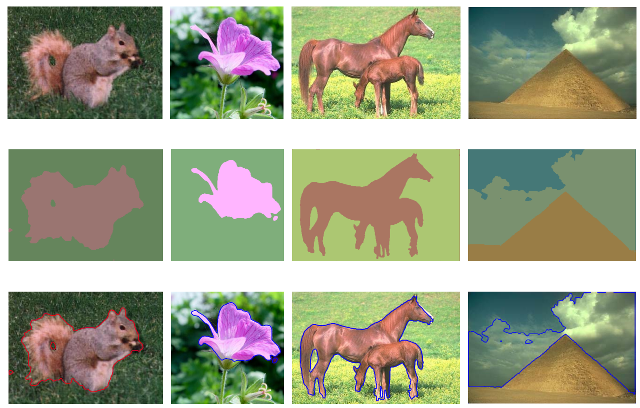
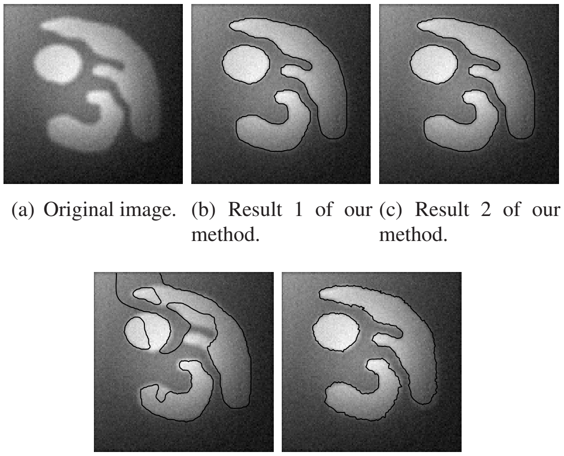


Figure 3. Application of our method for image segmentation. Column 1: The input images; Columns 2 and 3 show the segmentation results in the forms of regions and boundaries, respectively.

implements the Mumford-Shah model using a level set approach. We have implemented the PS model for gray level images in a two-phase formulation. For a fair comparison, we chose a synthetic gray level image in Fig. 4 that can be segmented by the PS model in a two-phase formulation. For this image, we applied our model with the membership regularization term. We used different initializations to test the robustness of the two methods. The results show that our method is robust to initialization. This can be seen from the results in Figs. 4(b) and 4(c) for two different initializations. The values of the initial membership functions are generated as random numbers for the first initialization (for the result in Fig. 4(b)), and a circular region is used to define the membership functions for the second initialization (for the result in Fig. 4(c)). It is clear that these two results are almost the same, which demonstrates the robustness of our method to initialization. The PS model often produces quite different results for even slightly different initializations in our implementation. The results of the PS model with two



(a) Original image. (b) Result 1 of our method. (c) Result 2 of our method.

(d) Result 1 of the PS model. (e) Result 2 of the PS model.

Figure 4. Comparison of our method with the piecewise smooth model.

initializations are shown in Figs. 4(d) and 4(e). The initialization of the level set functions in the PS model are obtained by two similar circular regions of different sizes. We also compared the CPU times of the two methods with the same circular region for the definition of the initial membership function in our method (for the result in Fig. 4(c))) and the level set function in the PS model (for the result in Fig. 4(e)). The CPU timed consumed by our method and the PS model are 7.83 and 141.97 seconds respectively. These CPU times are recorded in running our Matlab programs on a Lenovo ThinkPad notebook with Intel (R) Core (TM)2 Duo CPU, 2.40 GHz, 2 GB RAM, with Matlab 7.4 on Windows Vista. This experiment demonstrates the advantage of our method in terms robustness and computational efficiency over the PS model.

6. Conclusion

We have presented a novel method for image segmentation with simultaneous estimation of illumination and reflectance images. The proposed method is based on the composition of an observed scene image as the product of an illumination component and a reflectance component, known as intrinsic images. We define an energy functional in terms of an illumination image, the membership functions of the regions, and the corresponding reflectance constants of the regions in the scene. This energy is convex in each of its variables. By minimizing the energy, image segmentation result is obtained in the form of the membership functions of the regions. Meanwhile, the illumination and reflectance components of the observed image are estimated simultaneously as the result of energy minimization. With illumination taken into account, the proposed method is able to segment images with non-uniform intensities caused by spatial variations in illumination. Comparisons with the state-of-the-art piecewise smooth model demonstrate the superior performance of our method.

References

[1] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. In A. Hanson and E. Riseman, editors, *Computer Vision Systems*, pages 3–26. Academic Press, 1978.

[2] A. Blake. Boundary conditions for lightness computation in Mondrian world. *Computer Vision, Graphics and Image Processing*, 32(3):314–327, December 1985.

[3] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *J. Math. Imaging Vis.*, 28(2):151–167, 2007.

[4] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vis.*, 20(1-2):89–97, January 2004.

[5] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Imag. Proc.*, 10(2):266–277, February 2001.

[6] B. Horn. Determining lightness from an image. *Comput. Graphics, Image Processing*, 3(4):277–299, December 1974.

[7] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel. A variational framework for retinex. *Int'l J. Comp. Vis.*, 52(1):7–23, April 2003.

[8] E. Land and J. McCann. Lightness and retinex theory. *J. Opt. Soc. of America*, 61(1):1–11, 1971.

[9] Y. Matsushita, K. Nishino, K. Ikeuchi, and M. Sakauchi. Illumination normalization with time-dependent intrinsic images for video surveillance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(10):1336–1347, October 2004.

[10] S. E. Mila Nikolova and T. F. Chan. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM J. Appl. Math.*, 66(5):1632–1648, 2006.

[11] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[12] M. Tappen, E. Adelson, and W. Freeman. Estimating intrinsic component images using non-linear regression. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 1992–1999, 2006.

[13] L. Vese and T. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *Int'l J. Comp. Vis.*, 50(3):271–293, December 2002.

[14] Y. Weiss. Deriving intrinsic images from image sequences. In *Proceedings of the Eighth International Conference on Computer Vision (ICCV)*, volume II, pages 68–75, 2001.

[15] S.-C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(9):884–900, September 1996.