J. Vis. Commun. Image R. 30 (2015) 328–342

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Image segmentation via image decomposition and fuzzy region competition ☆

Yafeng Li

Department of Computer Science, Baoji University of Arts and Science, Baoji, Shaanxi 721016, China

ARTICLE INFO

Article history:
Received 16 June 2014
Accepted 16 April 2015
Available online 21 May 2015

Keywords:
Image segmentation
Image decomposition
Variational model
Wavelet
Dictionary learning
Sparse representation
Morphological component analysis
Variational optimization

ABSTRACT

Taking into account the morphological diversity of images, this paper presents a novel multiphase image segmentation method that combines image decomposition and fuzzy region competition into a unified model. To efficiently solve the minimization of the energy functional, we design an optimal iteration algorithm which integrates a modified cartoon-texture dictionary learning algorithm and wavelet shrinkage. Compared with the classical fuzzy region competition method, the proposed method not only improves the overall segmentation results, but also has more strong robustness. A series of experimental results demonstrate the applicability and effectiveness of the proposed method.

1. Introduction

Image segmentation is an important task for image analysis and pattern recognition. The goal of image segmentation is to partition an image into a set of disjoint sub-regions with uniform and homogeneous attributes. Many approaches have been proposed in recent years, among which variational method [1–8] is one of the most successful methods. In the variational framework, the level set (LS) method [2,3] and the fuzzy region competition (FRC) method [4–8] are widely used. The LS method uses level set functions to indicate different regions. The advantage of using LS functions to indicate the regions is that it automatically avoids the problems of overlap and vacuum. The main drawbacks of LS method are the slow convergence and initialization sensitiveness. Different from LS method, FRC method uses overlapping different fuzzy membership functions to indicate different regions. Due to the flexibility of fuzzy membership functions, more and more authors focus on fuzzy-based segmentation methods. In fact, Li et al. [5] have presented a multiphase FRC model. This model tries to find a piecewise constant function which approximates the observed image and uses total variation (TV) regularization on the fuzzy membership functions to prohibit the excessive length of the boundaries on regions. Later, Li et al. [6] use the nonparametric kernel density estimation method to segment texture images. Based on a nonconvex regularizer, Han et al. [7] propose a variational model for the fuzzy multiphase image segmentation. Choy et al. [8] present a multiphase fuzzy image segmentation model by taking into account both spatial and frequency information in data fidelity term. These works have shown very promising results.

This paper focuses on image segmentation using FRC method, and specifically explores the work of Li et al. [5]. The model in [5] assumes that image is piecewise smooth, therefore a piecewise constant function is used to approximate the given image. However, texture images or noisy images do not rigidly comply with the assumption of the model in [5]. So, the segmentation accuracy may be degraded in the case when image includes texture component or noise, specifically strong noise. The limitation motivates us to develop a more robust segmentation method. Texture is one of the main difficulties faced to a segmentation method. Note that, for texture image segmentation, many methods are based on the statistical or texture feature domain analysis [9,10]. Most segmentation algorithms cannot obtain texture features directly during segmentation. The existing segmentation methods lack a good representation of images including both their texture components and structure components. This paper presents a new method which can obtain both texture features and structure features represented by the learned dictionaries during segmentation. Different from image segmentation, image decomposition aims to separate the different features of image (different layers). The basic problem of image decomposition is cartoon-texture decomposition [11–14]. Because real images usually have two layers, namely, cartoon (the

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

E-mail address: liyafeng770729@126.com

http://dx.doi.org/10.1016/j.jvcir.2015.04.017
1047-3203/

piecewise smooth part of the image) and texture (the oscillating pattern part of the image), image decomposition has been successfully used in image denoising [14], image inpainting [15], texture discrimination [16] and blind source separation [17].

The major contribution of this work is that the proposed method unifies image segmentation and image decomposition. Taking into account the morphological diversity of images, we introduce image decomposition energy functional to the proposed segmentation model. The proposed method combines image decomposition and fuzzy region competition into a unified model. Due to the introducing of image decomposition term, the proposed method largely reduce the influence of illumination non-homogeneity and "texture noise" in the image. Hence, the proposed method not only improves the overall segmentation results, but also has more strong robustness.

The remainder of this paper is organized as follows. Some related works are reviewed in Section 2. The proposed model and the numerical algorithm are presented in Section 3. Comparative experiments are shown in Section 4. Finally, the conclusions are given in Section 5.

2. Related works

In this section, we briefly review the main frameworks of the multiphase FRC model in [5] and image decomposition model with learned dictionaries in [13].

2.1. The multiphase FRC image segmentation model

We first review the basic framework of the multiphase FRC model in [5]. Let Ω be the image domain and f be a given gray level image. The region number N is assumed to be known. The goal of image segmentation is to partition Ω into N disjoint sub-regions $(\Omega_i)_{i=1,2,\ldots,N}$ with uniform and homogeneous attributes. The multiphase FRC model based on the piecewise constant function in [5] is equivalent to the following minimization problem

$$\min_{0\leqslant I_i\leqslant 1}\left\{\sum_{i=1}^{N}\int_{\Omega}|\nabla I_i(x)|\mathrm{d}x+\lambda\sum_{i=1}^{N}\int_{\Omega}r_i^2(x)I_i(x)\mathrm{d}x+\frac{\gamma}{2}\int_{\Omega}\left(\sum_{i=1}^{N}I_i(x)-1\right)^2\mathrm{d}x\right\},\tag{1}$$

where $I_i(i=1,2,\ldots,N)$ in $[0,1]$ are fuzzy membership functions which indicate the regions $\Omega_i(i=1,2,\ldots,N)$. $\int_\Omega|\nabla I_i(x)|\mathrm{d}x$ is the bounded variation (BV) seminorm of I_i. It is also referred to as the total variation of I_i [18]. The denotation ∇ is the classical gradient operator. λ and γ are fixed positive parameters. The first term is a regularization term which characterizes some features of the desired solution $\hat{I}_i$. The second term is a data fidelity term which is defined to evaluate the performance of the label assignment at each partition Ω_i. A popular form $r_i(x)=f(x)-c_i$ is often used because the value c_i is easily updated in minimizing functional. In fact, the parameter c_i is the mean intensity value of f on Ω_i. The local version of the data fidelity term in the model (1) is also studied in the [5]. The third term is a quadratic penalty term which ensures the constraint $\sum_{i=1}^N I_i\approx 1$. There are two main drawbacks in the model (1). (i) The model assumes that the given image f is a piecewise smooth image. Hence, the piecewise constant function is used to approximate the original image. For the reason that the piecewise smoothing assumption is made in the model (1), the model can generate relatively good segmentation results when the images with simple structure is segmented. However, the model (1) is not quite suitable to segment texture images. (ii) The model (1) uses TV regularization on the fuzzy membership functions. TV regularization is well capable of preserving edges of uniform but it causes oversmoothing edges of disjoint regions, and may even destroy small scale structures with high curvature edges.

2.2. Image decomposition model with learned dictionaries

Image decomposition aims to separate different components in images (different layers of image). Real images usually have two layers, namely, cartoon (the piecewise smooth part of the image) and texture (the oscillating pattern part of the image). The image cartoon-texture decomposition problem can be modeled as $f=u+v$, where f is the given image, u is cartoon part, and v is oscillating texture or noise part. On the one hand, Meyer [11] constructed the theory of cartoon-texture decomposition problem. He proposed that cartoon component should be modeled by using total variation norm, and oscillating texture or noise part should be modeled by using the G-norm in the special space of harmonic analysis. But Meyer's variational model cannot be solved directly because of the existence of the weaker norm. Thus, a lot of authors begin to study the practical methods of Meyer's model [14,16,19]. On the other hand, Starck et al. [12] presented a sparse representation-based image decomposition method called morphological component analysis (MCA). The different components have sparse approximations under different dictionaries such as analytic dictionaries [12,20] and learned dictionaries [21–24]. Since analytic dictionaries lack the adaptivity, [13] has showed that learned dictionaries-based method has better performance than analytic dictionaries-based method for image decomposition. The authors in [13] propose an image decomposition model with the learned dictionaries as follows

$$\begin{aligned}&\left\{\hat{u},\hat{v},\widehat{D}_1,\widehat{D}_2,\hat{\alpha}_i,\hat{\beta}_i\right\}\\&=\operatorname*{argmin}_{u_1,u_2,D_1,D_2,\alpha_i,\beta_i}\left[\begin{array}{l}F(u,v,D_1,D_2,\alpha,\beta)\\=\tau\|f-u-v\|_2^2+\sum_{i\in P}\left(\|R_iu-D_1\alpha_i\|_2^2+\mu_i\|\alpha_i\|_0\right)\\+\sum_{i\in P}\left(\|R_iv-D_2\beta_i\|_2^2+\eta_i\|\beta_i\|_0\right)\end{array}\right],\\&s.t.\ D_1\in\mathfrak{I}_1,D_2\in\mathfrak{I}_2\end{aligned}\tag{2}$$

The first term is the log-likelihood global force. In the other two terms, the dictionaries $\boldsymbol{D}_1=[\boldsymbol{d}_1,\boldsymbol{d}_2,\ldots,\boldsymbol{d}_k]$ and $\boldsymbol{D}_2=[\boldsymbol{s}_1,\boldsymbol{s}_2,\ldots,\boldsymbol{s}_k]$ are used to respectively represent cartoon component u and texture component v under the sparse constraints l_0 quasi-norm. $\{\alpha_i\}_{i\in P},\{\beta_i\}_{i\in P}$ are sparse representation coefficients. P denotes the set of location indices. The operator R_i denotes extracting patch from location i in the cartoon part u and texture part v. λ,μ_i,η_i are trade-off parameters. $\mathfrak{I}_1$ and $\mathfrak{I}_2$ are two incoherence dictionary sets,

$$\mathfrak{I}_1=\{\boldsymbol{D}_1\in\mathbb{R}^{n\times k},\|\boldsymbol{d}_j\|\leqslant 1,j=1,\ldots,k\},$$
$$\mathfrak{I}_2=\left\{\boldsymbol{D}_2\in\mathbb{R}^{n\times k},\|\boldsymbol{s}_j\|\leqslant 1,\sum_{i=1}^{n}s_j[i]=0,j=1,\ldots,k\right\}$$

An alternate dictionary learning (A-DL) algorithm is proposed. Numerical experiments show that the learned dictionaries by the A-DL algorithm can describe the different components of image effectively and leads to high quality image decomposition performance. For more details, the reader can refer to Appendix A.

3. The proposed model and algorithm

3.1. The proposed model

In order to overcome the drawbacks of the model (1), this paper takes into account the morphological diversity of image, and