## WAVELET FRAME BASED MULTI-PHASE IMAGE SEGMENTATION

## CHENG TAI\*, XIAOQUN ZHANG<sup> $\dagger$ </sup>, AND ZUOWEI SHEN<sup> $\ddagger$ </sup>

Abstract. Wavelet frames have been successfully applied in various image restoration problems, such as denoising, inpainting, deblurring, etc. However, they are rarely used in geometric applications, except for recent work of [22, 23]. Motivated by the theoretical connection between wavelet frame based and total variation based image restoration models recently established in [7] we propose here a convex multi-phase segmentation model based on wavelet frame transform. The proposed model allows to automatically identify complex tubular structures, including blood vessels, leaf vein system, etc. Numerical results show that our method can extract many more details than existing variational methods especially when the image contains different scales of structures. The proposed method can be parallelized easily and its efficiency is further improved by a GPU implementation. In addition, we analyze the connection between solutions of the convexified model and the original binary constrained one.

 ${\bf Key}$  words. image segmentation, wavelet frame, multi-phase partitioning, multiscale, convex relaxation

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. Multi-phase image segmentation, or multi-phase labeling is the process of partitioning an image into multiple regions with respect to specific goals and applications. It is a basic but very important image analysis task that has been extensively investigated for many years. Among all the models, regularization based variational models have proven to be especially successful. Variational models started with the classic work by Mumford and Shah [38] and active contour models [31, 35]. Later, Chan-Vase active contour model [18] and its variants based on level sets and total variation [42] were proposed to improve earlier results in terms of both segmentation accuracy and computation efficiency, see [17, 18, 32, 33]. However, the quality of Chan-Vese model relies on initializations due to the non-convexity of the model. Meanwhile, it is well known that Mumford-Shah model is a special case of classical Pott's model in discrete setting. The general Pott's model consists of solving the image segmentation problem by minimizing a sum of the lengths of the boundaries of the regions and data fidelity. It is well known that solving Pott's model is NP hard and can not be solved in a polynomial time for multi-phase cases. Many kinds of graph-cut based models such as alpha expansion, alpha-beta swap [4] have been developed to approximate global minimization solution. Recently, convexified models by relaxation of binary partition in both discrete and continuous setting [16, 5, 6, 2, 48]were proposed to improve the robustness of segmentation as well as computational efficiency. A detailed study on the convex model based on total variation is recently present in [12].

On the other hand, the theory of tight wavelet frames, also called framelets, has been extensively studied in the past two decades with many successful applications in image processing, including image denoising, image deblurring, inpainting, etc, see

1

<sup>\*</sup>The program in applied and computational mathematics, Princeton University(chengt@princeton.edu).

<sup>&</sup>lt;sup>†</sup>Department of mathematics, MOE-LSC and Institute of natural sciences, Shanghai Jiao Tong University(xqzhang@sjtu.edu.cn). X. Zhang was partially supported by the NNSFC (Grant no. 11101277) and the Shanghai Pujiang Talent Program (Grant no. 11PJ1405900)

<sup>&</sup>lt;sup>‡</sup>Department of mathematics, National University of Singapore(matzuows@nus.edu.sg)

[45, 27, 8, 24, 7] and references therein. Examples of tight frames include translation invariant wavelets, wavelet frame and curvelets, etc [21, 40, 41, 10]. Unlike orthogonal wavelets, tight frames give redundant representations of signals. The redundancy of tight frames often provides the flexibility that is a desired property in various applications. Despite the success of tight frames and wavelet frame in many applications in image processing, there are few geometric applications such as image segmentation. Motivated by recent theoretical progress in [7] that establishes the fundamental connection between wavelet frames based and total variation based approaches for image restoration, we intend to continue the attempt of [22] and further explore the application of wavelet frames for image segmentation. Some previous attempts to do so have been made in [22] and [9]. In [22], the authors proposed a two-phase segmentation model utilizing wavelet frame. In [9], the authors also proposed an iterative procedure with thresholding on wavelet frame coefficients to segment tubular structures. Both papers use numerical results to show that wavelet frame based models are superior to existing variational segmentation models, especially for fine structures. Both models deal with two-phase segmentation. There are also some previous work on texture classification using wavelet frames, see [46, 1]. But to the best of our knowledge, utilizing the property of sparse approximation of tight frames and convex approaches for multi-phase segmentation has not been well studied in the literature. This paper aims to fill this gap.

We derive a multi-phase segmentation model based on wavelet frame and convexified segmentation model. The proposed model is applied to automatically identify complex tubular structures, including blood vessels in magnetic resonance angiography images and leaf vein systems. These kinds of images are challenging due to intensity inhomogeneity, intersection of different scales of structures and the presence of noise. The early results from [22] and the nature of the wavelet frame based approach indicate that the wavelet frame performs efficiently and specially well when the intensity is not homogenous, which is another motivation of our current adventure. The quality of segmentation is crucial for further structure analysis. One major difficulty shared by the two kinds of images is that they contain different scale of structures and existing segmentation algorithms may not be able to get satisfactory results. For existing algorithms for identifying blood vessels, interesting reader should consult [29, 28, 36] for details. Wavelet frames, constructed from multi-resolution analysis (MRA), adapt to different scales naturally, and therefore are suitable to these applications. The advantage of using wavelet frame to do segmentation is partially proved in the two-phase case in [22, 9]. More generally, the tubular structure often belongs to one of the many regions in a given image, therefore, a multi-phase segmentation model is desirable. The proposed model can segment the given image into multiple regions and utilize some properties of wavelet frame such as natural multi-scale description of structures as well as sparse approximation of piecewise smooth images.

Like other variational models, the proposed model yields a minimization problem. There are a variety of frameworks and algorithms that aim to solve sparse optimization problems, for example [47, 30, 8, 25, 12, 50, 49]. In particular, we apply a first order primal-dual framework [25, 12] to build our algorithm due to its efficiency and simplicities. Details will be given in later sections. We also point out that unlike usual image restoration model where low frequencies coefficients is never penalized, using of low frequencies encourages binary solutions and also yield faster numerical convergence. Finally, numerical results show that the proposed model extracts many more details than total variation based models especially when the input image contains



Fig. 1.1: Fine structure segmentation

different scales of structures and singularities in low contrast setting.

The rest of this paper is organized as follows: in the rest of this section, we present variational image segmentation models and our motivations of using wavelet frame, and then review some concepts for wavelet frame and the basic framework of primal-dual algorithms. In section 2, we present the formulation of wavelet frame based multi-phase segmentation model and give the detail of the algorithm to solve the minimization problem. We also analyze the connections between the solutions of the convexified and the original binary models. In section 3, we give some numerical results and compare our results with some existing models. We also give remarks on the role of low frequency. Sections 4 concludes the paper.

**1.1. Variational segmentation model and convex relaxation.** Given an image I(x) defined on an image domain  $\Omega \subset \mathbb{R}^d$  for d = 2, 3, the image segmentation problem is to find a partition of  $\Omega$  into K disjoint subdomains  $\{\Omega_k\}_{k=1}^K$ , that is:

$$\Omega = \bigcup_{k=1}^{K} \Omega_k, \ \Omega_k \bigcap \Omega_j = \emptyset, \text{ if } k \neq j.$$

Mumford-Shah [38] proposed minimizing the interface between the partitions with a piecewise constant variational model:

$$\min_{\Omega_k, c_k} \left\{ \sum_{k=1}^K |\partial \Omega_k| + \frac{\lambda}{2} \sum_{k=1}^K \int_{\Omega_k} |I(x) - c_k|^2 dx \right\},\tag{1.1}$$

where  $c_k \in \mathbb{R}$  for  $k = 1, \dots, K$  is the mean value in each region  $\Omega_k$ . The parameter  $\lambda > 0$  is used to balance the data fitting term and the total length of interfaces. It is well known that this model is a special case of classical Pott's model where the fidelity term is given in a more general setting. The discreet Pott's model is NP hard and to simplify the task, it is often assumed that the mean value  $c_i$  of each region is known and fixed. The model (1.1) is thus used to look for a smooth and tight boundary between regions.

By introducing the labelling function  $u_k$  of the disjoint subdomains  $\Omega_k$ 

$$u_k(x) = \begin{cases} 1 & \text{if } x \in \Omega_k \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = 1, \cdots, K.$$
 (1.2)

we can formulate the model (1.1) in the generic form

$$\min_{\boldsymbol{u}} \left\{ \sum_{k=1}^{K} |\partial \Omega_k| + \sum_{k=1}^{K} \int_{\Omega} u_k(x) f_k(x) dx \right\}$$
s.t.  $u_k(x) = \{0, 1\}, \sum_{k=1}^{K} u_k(x) = 1, \quad \forall x \in \Omega$ 
(1.3)

where  $f_k(x) = \frac{\lambda}{2} |I(x) - c_k|^2$ .

The interface lengths can be further convert to total variation thanks to the coarea formula [17, 13]

$$\sum_{k=1}^{K} |\partial \Omega_k| = \sum_{k=1}^{K} \int_{\Omega} |\nabla u_k(x)| dx$$
(1.4)

This transform allows us to develop efficient algorithms based on the well studied total variation minimization [42, 16, 13]. However, the above model is nonconvex due to the binary constraint of the  $u_k$ . Generally, a convex relaxation is made by relaxing the binary constraint over the interval [0, 1]. By denoting  $\boldsymbol{u} = (u_1, \dots, u_K)$ ,  $\boldsymbol{f} = (f_1, \dots, f_K)$  and  $J(\boldsymbol{u}) = \sum_{k=1}^K \int_{\Omega} |\nabla u_k(x)| dx$ , the convex minimization problem is formulated as:

$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}\in\boldsymbol{S}} \left\{ J(\boldsymbol{u}) + \langle \boldsymbol{u}, \boldsymbol{f} \rangle \right\}$$
(1.5)

where S is the simplex constraint at each pixel and defined as

$$\boldsymbol{S} = \left\{ \boldsymbol{u}: \Omega \to \mathbb{R}^K : u_k(x) \in [0, 1], \text{ for } 1 \le k \le K; \sum_{k=1}^K u_k(x) = 1, \forall x \in \Omega \right\}$$
(1.6)

and the inner product  $\langle \boldsymbol{u}, \boldsymbol{f} \rangle = \sum_{k=1}^{K} \int_{\Omega} u_k(x) f_k(x) dx$  is understood in usual sense. If the minimizer of (1.5) happens to be binary everywhere, then it is also a global minimizer of the original problem (1.3). On the other hand, a global minimizer of (1.5) might not be binary even when the solution of (1.3) is unique. Generally, a final thresholding step is taken to get a binary solution

$$u_k^*(x) = \begin{cases} 1 & \text{if } u_k^*(x) = \max\{u_1^*(x), u_2^*(x), \dots, u_K^*(x)\} \\ 0 & \text{otherwise} \end{cases}$$
(1.7)

If the maximizer is not unique, the maximizer with smallest subscript is generally used as a convention.

This region based model has also been combined with an edge based approach to generalize the above model, such as in [43, 5, 3]. By introducing an edge indicator  $g(x) \ge 0$ , the model (1.5) is extended by setting

$$J(\boldsymbol{u}) = \sum_{k=1}^{K} \int_{\Omega} g(x) |\nabla u_k(x)| dx$$
(1.8)

The edge indicator function takes small values at locations with large gradient and big values for smooth region. For a given image I(x), a usual choice is

$$g(x) = \frac{1}{1 + \sigma \|\nabla \tilde{I}(x)\|^2}$$
(1.9)

where I is a smoothed version of I and  $\sigma > 0$  is a positive number. Note that if g(x) is identically 1, it reduces to the model (1.5).

The above total variation based methods have been proved to be very efficient for segmenting piecewise constant types of cartoon-like images, while it may not so efficient for complex tubular or vascular structures such as blood vessel, vein systems in medical images. Recently, connections between total variation and wavelet frame based image restoration models are established in [7]. One of main results of this work shows that total variation based image restoration model can be viewed as the limit of a wavelet frame based model when the resolution goes to infinite. The analysis therein provides geometric interpretations to wavelet frames approach as well as its solutions. The successes of wavelet frame based images motivate us to further investigate the application of wavelet frame based method on image segmentation with multiple phases. Based on the theoretical connection to continuous differential operator and wavelet frame filters given in [7], wavelet frame based approaches can adaptively choose proper filter of different vanishing moment according to the order of the singularity of the underlying images. This is particularly useful for segmentation since the main goal of segmentation is to capture singularities of different order and scale which represent edges in images. These observations will enable us to design new multi-phase image segmentation methods based on wavelet frame, especially for images with multilevel fine structures in very low contrast setting.

**1.2. MRA based wavelet frames.** In this section, we briefly introduce the concepts of tight frames and wavelet frames. Interested readers should refer to [20, 40, 41, 21, 24] for the theories and applications of wavelet frames.

A countable set  $X \subset L_2(\mathbb{R}^d)$  is called a tight frame of  $L_2(\mathbb{R}^d)$  if

$$f = \sum_{\psi \in X} \langle f, \psi \rangle, \quad \forall f \in L_2(\mathbb{R}^d),$$

where  $\langle , \rangle$  is the inner product of  $L_2(\mathbb{R}^d)$ 

For a given set  $\Psi = \{\psi_1, \ldots, \psi_r\} \subset L_2(\mathbb{R}^d)$ , the affine system is defined by the collection of the dilations and shifts of  $\Psi$  as

$$X(\Psi) = \left\{ \psi_{l,i,\boldsymbol{k}} : l \in \mathbb{Z}, 1 \le i \le r, \boldsymbol{k} \in \mathbb{Z}^d \right\} \text{with } \psi_{l,i,\boldsymbol{k}} = 2^{l/2} \psi_i (2^l \cdot -\boldsymbol{k}).$$
(1.10)

When  $X(\Psi)$  forms a tight frame of  $L_2(\mathbb{R}^d)$ , it is called a tight wavelet frame, and  $\psi_l, l = 1, \ldots, r$  are called (tight) wavelet frame or framelets.

To construct a set of framelets, one usually starts from a compactly supported refinable function (also called scaling function) that generates a multi-resolution analysis (MRA) space of  $L_2(\mathbb{R}^d)$ , satisfying

$$\widehat{\phi}(2\cdot) = \widehat{h}_0 \widehat{\phi}$$

for some  $h_0$ . Here  $\hat{\phi}$  is the Fourier transform of  $\phi$ , and  $\hat{h}_0$  is the Fourier series of  $h_0$ . For a given compactly supported refinable function, a tight framelet system is constructed by finding a finite set  $\Psi$  that can be represented in the Fourier domain as

$$\widehat{\psi}_l(2\cdot) = \widehat{h}_l \widehat{\phi}$$

for some  $2\pi$ -periodic function  $\hat{h}_l$ . The unitary extension principle (UEP)[40, 41] shows that the system in (1.10) generated by  $\Psi$  forms a tight frame in  $L_2(\mathbb{R})$  provided that the masks of  $\hat{h}_l$  for  $l = 0, 1, \ldots, r$  satisfy

$$\sum_{i=0}^{r} |\hat{h}_{i}(\xi)|^{2} = 1 \quad \text{and} \quad \sum_{i=0}^{r} \hat{h}_{i}(\xi) \overline{\hat{h}_{i}(\xi + \pi)} = 0$$
(1.11)

for almost all  $\xi$  in  $\mathbb{R}$ . Here  $h_0$  corresponds to a low pass filter, while  $\{h_i : i = 1, 2, ..., r\}$  must correspond to high-pass filters by UEP (1.11). The sequences of Fourier coefficients of  $\{h_i : i = 1, 2, ..., r\}$  are called framelet masks.

In our implementation, we adopt two kinds of wavelet frame. One is constructed from a piecewise linear B-spline. The filter banks coefficients are

$$h_0 = \frac{1}{4}[1,2,1], \quad h_1 = \frac{\sqrt{2}}{4}[1,0,-1], \quad h_2 = \frac{1}{4}[-1,2,-1].$$

The other is also constructed from B-splines but with higher vanishing moments. The filter banks are

$$h_0 = \begin{bmatrix} \frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16} \end{bmatrix}, \quad h_1 = \begin{bmatrix} \frac{1}{16}, -\frac{1}{4}, \frac{3}{8}, -\frac{1}{4}, \frac{1}{16} \end{bmatrix}, \quad h_2 = \begin{bmatrix} -\frac{1}{8}, \frac{1}{4}, 0, -\frac{1}{4}, \frac{1}{8} \end{bmatrix}$$
$$h_3 = \begin{bmatrix} \frac{\sqrt{6}}{16}, 0, -\frac{\sqrt{6}}{8}, 0, \frac{\sqrt{6}}{16} \end{bmatrix}, \quad h_4 = \begin{bmatrix} -\frac{1}{8}, -\frac{1}{4}, 0, \frac{1}{4}, \frac{1}{8} \end{bmatrix}.$$

The framelet transform can be obtained by convolving the signal with these filter bank coefficients. With a one-dimensional framelet system for  $L_2(\mathbb{R}^d)$ , the *d*-dimensional framelet system for  $L_2(\mathbb{R}^d)$  can be easily constructed by tensor products of one dimensional wavelet frame.

**1.3. Primal-Dual algorithm.** In this section, we briefly review a class of primal dual algorithms that we will apply later for solving our proposed model. The general optimization problem that we consider takes the following form :

$$\min_{u \in X} \left\{ F(Bu) + G(u) \right\} \tag{1.12}$$

where X is convex set in  $\mathbb{R}^d$ ,  $G: X \mapsto [0, +\infty)$  and  $F: X \mapsto [0, +\infty)$  are two proper, convex, lower-semicontinuous (l.s.c) functions and the map  $B: X \mapsto Y$  is a continuous linear operator with induced norm:

$$||B|| = \max\{||Bu||_Y : u \in X \text{ with } ||u||_X \le 1\}.$$

Generally, the convex functional F is not differentiable, e.g.  $\ell^1$  norm type, thus making the optimization difficult to solve. There is considerable interest in finding efficient algorithms to solve this optimization problem recently due to its applications arising in imaging sciences, such as total variation based image restoration, inverse problem regularization, low rank matrix restoration, etc. In [30], a splitting based method was proposed to solve this problem efficiently for total variation case. In [50], a primal dual hybrid method (PDHG) method was proposed to solve total variation regularization problem. In [25, 12] more general algorithm framework has been proposed for this type of problem.

Using Fenchel-Legendre transform, we obtain the generic saddle-point problem:

$$\min_{u \in X} \max_{p \in Y} \left\{ \langle Bu, p \rangle + G(u) - F^*(p) \right\}, \tag{1.13}$$

where  $F^*$  is the convex conjugate of the l.s.c. function F. The corresponding dual problem is written as

$$\max_{p \in Y} \left\{ -(G^*(-B^*p) + F^*(p)) \right\}.$$
(1.14)

where  $B^*$  is the adjoint operator of B.

We assume that these problems have at least one solution  $(\hat{u}, \hat{p}) \in X \times Y$ , which satisfies:

$$B\hat{u} \in \partial F^*(\hat{p})$$

$$-(B^*\hat{p}) \in \partial G(\hat{u}),$$

where  $\partial F^*$  and  $\partial G$  are the subgradients of the convex function  $F^*$  and G. The resolvent operator (also known as the proximity operator) is defined by

$$u = (I + \tau \partial G)^{-1}(v) = \arg\min_{v \in X} \left\{ \frac{\|u - v\|^2}{2\tau} + G(u) \right\}.$$
 (1.15)

Assume F and G are simple in the sense that this resolvent has a closed form. With these notations, the primal-dual algorithm proposed in [25, 12] is given as follows:

 $\begin{array}{l} \hline \textbf{Algorithm 1 Modified PDHG algorithm [25, 12].} \\ \hline \textbf{Require: Choose } \tau, \sigma > 0, (x^0, y^0) \in X \times Y \text{ and set } \bar{x}^0 = x^0 \\ 1: \ \textbf{for } n \geq 0, \text{ update } p^{n+1}, u^{n+1}, \bar{u}^{n+1} \ \textbf{do} \\ 2: & \left\{ \begin{array}{l} p^{n+1} = (I + \sigma \partial F^*)^{-1} (p^n + \sigma B \bar{u}^n) \\ u^{n+1} = (I + \tau \partial G)^{-1} (u^n - \tau B^* p^{n+1}) \\ \bar{u}^{n+1} = 2u^{n+1} - u^n \end{array} \right. \\ 3: \ \textbf{end for.} \end{array}$ 

As shown in [25, 12], if  $\tau\sigma\beta^2 < 1$  where  $\beta = ||B||_2$ , the sequence defined by Algorithm 1 converges and the primal dual gap has the convergence rate  $\mathcal{O}(1/N)$  in ergodic sense.

## 2. Wavelet frame based segmentation: model and algorithm.

**2.1. Notations.** In this section, we consider the discrete setting of wavelet frame transform. For simplicity, we still denote the pixel/voxel index set as  $x \in \Omega$ , where  $\Omega$  is the image domain in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ . Let  $n = |\Omega|$  be the number of pixels of image domain,  $U = \mathbb{R}^n$  be the vector space of images that region indicator u(x) belongs to regardless of its dimension.

We use W to denote the fast tensor product framelet decomposition and denote an *L*-level framelet decomposition of u as

$$\boldsymbol{W}\boldsymbol{u} = \left\{ W_0\boldsymbol{u}; W_{l,\boldsymbol{i}}\boldsymbol{u}; \quad \text{for } \forall 0 \leq l \leq L-1; \ \forall \boldsymbol{i} \in \mathcal{I} \right\},\$$

where  $W_0$  denotes the lowest frequency at decomposition level L and

$$\mathcal{I} = \{ (i_1, \cdots, i_d), \quad \text{for } 0 \le i_1, \cdots, i_d \le r, (i_1, \cdots, i_d) \ne (0, \cdots, 0) \}$$
(2.1)

denotes the index set of all framelet bands, and l denotes the decomposition level. We use  $\boldsymbol{W}^{\top}$  to denote the fast reconstruction and by UEP we have  $\boldsymbol{W}^{\top}\boldsymbol{W} = I$ , i.e.  $u = \boldsymbol{W}^{\top}\boldsymbol{W}u$  for any image u.

We denote the set of wavelet frame coefficients as  $\mathbb{P}.$  For each element in  $p\in\mathbb{P},$  it can be expressed as

$$p = \left\{ p_0; p_{l, \boldsymbol{i}}, \ p_0 \in \mathbb{R}^n; p_{l, \boldsymbol{i}} \in \mathbb{R}^n \quad \text{for } \forall 0 \le l \le L - 1; \ \forall \boldsymbol{i} \in \mathcal{I} \right\}$$

Here  $p_0(x)$  denotes the low frequency coefficients vector, and  $p_{l,i}$  denotes the coefficients vector at filterband i and decomposition level l.

Furthermore, we introduce the discrete  $\ell_{1,2}$  semi-norm of a framlet-coefficient vector  $p \in \mathbb{P}$  without low frequency:

$$\|p\|_{1,2}^{0} = \left\|\sum_{l=0}^{L-1} \left(\sum_{\boldsymbol{i}\in\mathcal{I}} |p_{l,\boldsymbol{i}}|^{2}\right)^{1/2}\right\|_{1}$$
(2.2)

where  $|\cdot|^2$  and  $(\cdot)^{\frac{1}{2}}$  are componentwise operations and  $\|\cdot\|_1$  denotes the  $\ell_1$  norm in  $\mathbb{R}^n$ . The summation goes over all the *L* levels and the set of bands except the low frequencies. For this reason, this definition is a semi-norm. This is usually used for image restoration application, see [8, 24]. It is also used in the two-phase image segmentation in [22]. Here, we are also interested in the norm including the low frequency, defined as

$$\|p\|_{1,2} = \|p_0\|_1 + \|p\|_{1,2}^0$$
(2.3)

Finally, given image I(x) for  $x \in \Omega$ , we want to segment in K subregions  $\{\Omega_1, \dots, \Omega_K\}$ . We denote  $U = U^K$  as the product space of K region indicator functions  $u = (u_1, \dots, u_K)$  belongs to. Let  $P = \mathbb{P}^K$  be the corresponding product space of K- wavelet frame coefficients vector spaces. Thus for any  $u \in U$ , we denote  $Wu = (Wu_1, \dots, Wu_k, \dots, Wu_K)$  and  $Wu \in P$ .

**2.2.** Model. As we stated previously, total variation based image restoration model can be viewed as the limit of a wavelet frame based model when the resolution goes to infinite. The theoretical connection to continuous differential operator and wavelet frame filters given in [7] provides geometric interpretations to wavelet frames approach as well as its solutions. The successes of wavelet frame for image restoration motivate us to further investigate the application of wavelet frame to image segmentation with multiple phases.

We consider the general model (1.5). For the fidelity term, we first assume that we have an initial guess of the mean of each subregion  $c_k$  for  $k = 1, \dots, K$  and

$$f_k(x) = \frac{\lambda}{2} \|I(x) - c_k\|^{\beta}$$
 (2.4)

for  $\beta = 1, 2$ .

For the regularization term in (1.5), total variation can be considered as a special case of framelet coefficients  $\|\cdot\|_{1,2}^0$  norm defined in (2.3) when Harr filter is considered as shown in [7]. We consider wavelet frame regularization instead of total variation since wavelet frame based approaches can adaptively choose proper filter of different

vanishing moment according to the order of the singularity of the underlying images. This is particularly useful for segmentation since the main goal of segmentation is to capture singularities of different order and scale which represent edges in images. By introducing an edge function  $g \in \mathbb{R}^n$  with  $g(x) \geq 0$  for each  $x \in \Omega$  and generalizing the definition of the norm, we define the weighted tight frame  $\ell_{1,2}$  norm for a given image  $u_k(x)$ 

$$\|g \cdot \mathbf{W}u_k\|_{1,2} := \langle g, |W_0 u_k| \rangle + \left\langle g, \sum_{l=0}^{L-1} \left( \sum_{i \in \mathcal{I}} |W_{l,i} u_k|^2 \right)^{1/2} \right\rangle$$

$$= \|g \cdot W_0 u_k\|_1 + \|g \cdot \mathbf{W}u_k\|_{1,2}^0$$
(2.5)

where  $|\cdot|, |\cdot|^2$  and  $(\cdot)^{1/2}$  are componentwise for each  $x \in \Omega$ .

We point out here that unlike applications in image restorations, where the low frequency part of the framelet transform is not used, the low frequency is used and it has several additional advantages for segmentation application. Therefore, our proposed binary constrained wavelet frame based multi-phase segmentation model is formulated as

$$\min_{(u_1, \cdots, u_K)} \left\{ \sum_{k=1}^{K} \|g \cdot \boldsymbol{W} u_k\|_{1,2} + \sum_{k=1}^{K} \langle u_k, f_k \rangle \right\}$$
(2.6)  
s.t.  $u_k(x) \in \{0, 1\}, \quad \sum_{k=1}^{K} u_k(x) = 1, \quad \forall x \in \Omega$ 

For a given edge indicator function g(x),  $x \in \Omega$ , g(x) takes small values when the transition of the image intensity is sharp at  $x \in \Omega$ . It is used to slow down the evolution of the interface when it arrives at the boundaries. In our case, we may use the high frequency part to measure this transition, which is similar to the gradient in total variation based models. In our implementation, we set

$$g(x) = \frac{1}{1 + \sigma \sum_{l, \boldsymbol{i} \in \mathcal{I}} |(W_{l, \boldsymbol{i}} \tilde{I})(x)|^{\beta}}$$
(2.7)

for all  $x \in \Omega$ , where  $\sigma > 0$  is a given positive number and  $\tilde{I}(x)$  is a smoothed image of the input I(x). Other possible choice of edge indicator function can be also used.

The model is not convex due to the binary constraints on  $\boldsymbol{u}$ , which brings much numerical difficulty. As considered in several previous work, we relax the constraints so that each  $u_k$  is allowed to take values continuously from [0, 1]. Along with the relaxation, the constraint becomes  $\sum_k u_k(x) = 1$ . Finally, the convex relaxation model in vector form is written as

$$\min_{\boldsymbol{u}\in\boldsymbol{S}} E^{P}(\boldsymbol{u}) := \|g\cdot\boldsymbol{W}\boldsymbol{u}\|_{1,2} + \langle \boldsymbol{u},\boldsymbol{f}\rangle$$
(2.8)

where  $S \subset U$  is the simplex constraint defined in (1.6) and

$$\|g \cdot \boldsymbol{W}\boldsymbol{u}\|_{1,2} = \sum_{k=1}^{K} \|g \cdot \boldsymbol{W}u_k\|_{1,2} = \|g \cdot W_0\boldsymbol{u}\|_1 + \|g \cdot \boldsymbol{W}\boldsymbol{u}\|_{1,2}^0, \quad (2.9)$$

FRAMELET BASED IMAGE SEGMENTATION

$$\langle \boldsymbol{u}, \boldsymbol{f} \rangle = \sum_{k=1}^{K} \langle u_k, f_k \rangle$$
 (2.10)

**2.3. Algorithm.** In this subsection, we give the details of the algorithm for solving the optimization problem above. The main framework is based on a primal dual algorithm introduced in both [25] and [12], as described in Algorithm 1.

We can reformulate the proposed model (2.8) in the general form of (1.12) in Section 1.3. Let

$$F(\boldsymbol{W}\boldsymbol{u}) = \left\| g \cdot \boldsymbol{W}\boldsymbol{u} \right\|_{1,2}, \qquad G(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{f} \rangle + \delta_{\boldsymbol{S}}(\boldsymbol{u}),$$

where  $\delta_{\mathbf{S}}(u)$  is the characteristic function of convex set  $\mathbf{S}$  defined in (1.6), i.e.

$$\delta_{\boldsymbol{S}}(\boldsymbol{u}) = \begin{cases} 0 & \text{if } \boldsymbol{u} \in \boldsymbol{S} \\ \infty & \text{otherwise.} \end{cases}$$
(2.11)

For this general form, we have the primal-dual model as (1.13):

$$\min_{\boldsymbol{u}\in\boldsymbol{U}}\max_{\boldsymbol{p}\in\boldsymbol{P}}\left\{\langle \boldsymbol{W}\boldsymbol{u},\boldsymbol{p}\rangle+G(\boldsymbol{u})-F^{*}(\boldsymbol{p})\right\},$$
(2.12)

where  $\boldsymbol{p} = (p^{(1)}, \dots, p^{(K)})$  and each  $p^{(k)} \in \mathbb{P}$  corresponds to a dual variable of  $u_k$  in the tight wavelet frame coefficients space, i.e.

$$p^{(k)} = \left\{ p_0^{(k)}(x); p_{l,\boldsymbol{i}}^{(k)}(x), \text{ for } 0 \le l \le L - 1, \boldsymbol{i} \in \mathcal{I}, x \in \Omega \right\}$$
(2.13)

We need to compute the resolvent operator (1.15)  $(I + \partial F^*)^{-1}$  and  $(I + \partial G)^{-1}$  for applying Algorithm 1. In the following, we first compute the conjugate function  $F^*$ . For notational convenience, we first drop the index k. Let  $q = (q_0(x), q_{l,i}(x)) \in \mathbf{P}$ be a vector in the wavelet frame transform domain and F(q) denote the g- weighted  $\ell_{1,2}$  norm of q:

$$F(q) = \|g \cdot q\|_{1,2} = \langle g, |q_0| \rangle + \langle g, \sum_{l=0}^{L-1} \left( \sum_{i} |q_{l,i}|^2 \right)^{1/2} \rangle$$
(2.14)

By the definition of a conjugate function, we have

$$\begin{split} F^*(p) &= \max_{q} \left\{ \langle p, q \rangle - F(q) \right\} \\ &= \max_{q} \left\{ \langle q_0, p_0 \rangle - \langle g, \ |q_0| \rangle + \sum_{l=0}^{L-1} \left( \sum_{\boldsymbol{i}} \langle p_{l,\boldsymbol{i}}, q_{l,\boldsymbol{i}} \rangle - \langle g, \ \left( \sum_{\boldsymbol{i}} |q_{l,\boldsymbol{i}}|^2 \right)^{1/2} \rangle \right) \right\} \\ &= \left\{ \begin{array}{c} 0, \quad \text{if } |p_0(x)| \leq g(x) \text{ and } \max_{l} (\sum_{\boldsymbol{i} \in \mathcal{I}} \left| p_{l,\boldsymbol{i}}(x) \right|^2)^{1/2} \leq g(x) \quad \forall x \in \Omega \\ \infty, \quad \text{otherwise} \end{array} \right. \end{split}$$

Denote  $Y \subset P$  as

$$\boldsymbol{Y} = \left\{ p = \left( p_0(x), p_{l,\boldsymbol{i}}(x) \right) : |p_0(x)| \le g(x); \max_{0 \le l \le L-1} \left( \sum_{\boldsymbol{i} \in \mathcal{I}} \left| p_{l,\boldsymbol{i}}(x) \right|^2 \right)^{1/2} \le g(x) \quad \forall x \in \Omega \right\}$$

$$(2.15)$$

10

We can easily see that  $\boldsymbol{Y}$  is a convex set.

Now if we consider K regions, for a vector  $\boldsymbol{p} = (p^{(1)}, \dots, p^{(K)}) \in \boldsymbol{P}$ , we have  $F^*(\boldsymbol{p})$  as the characteristic function of convex set  $\boldsymbol{Y}$  for each  $p^{(k)}$ :

$$F^*(\boldsymbol{p}) = \begin{cases} 0, & \text{if } p^{(k)} \in \boldsymbol{Y}, k = 1, \cdots, K \\ \infty, & \text{otherwise} \end{cases}$$

For a given tight frame coefficients vector  $\boldsymbol{q} = (q^{(1)}, \cdots, q^{(K)})$ , the resolvent operator defined in (1.15) for  $F^*$  (2.16) is given by

$$(I + \sigma \partial F^*)(\boldsymbol{q}) = \arg \min_{\boldsymbol{p}} \left\{ \frac{\|\boldsymbol{p} - \boldsymbol{q}\|^2}{2\sigma} + F^*(\boldsymbol{p}) \right\}$$
$$= \arg \min_{\boldsymbol{p}^{(k)} \in \boldsymbol{Y}} \left\{ \frac{\|\boldsymbol{p} - \boldsymbol{q}\|^2}{2\sigma} \right\}$$
$$= \left( \prod_{\boldsymbol{Y}} (q^{(1)}), \cdots, \prod_{\boldsymbol{Y}} (q^{(K)}) \right)$$

where  $\prod_{\boldsymbol{Y}}$  denotes the projection operator onto the convex set  $\boldsymbol{Y}$  for each  $q^{(k)}$ . Let  $\tilde{q} = \prod_{\boldsymbol{Y}} (q)$ . We drop the index k for simplicity. For  $0 \leq l \leq L-1$ , high frequency coefficients  $\tilde{q}_{l,\boldsymbol{i}}, \boldsymbol{i} \in \mathcal{I}$  are given by

$$\tilde{q}_{l,\boldsymbol{i}}(x) = \begin{cases} q_{l,\boldsymbol{i}}(x), & \text{if } \left(\sum_{\boldsymbol{i}\in\mathcal{I}} |q_{l,\boldsymbol{i}}(x)|^2\right)^{1/2} \leq g(x) \\ \frac{g(x)q_{l,\boldsymbol{i}}(x)}{\left(\sum_{\boldsymbol{i}\in\mathcal{I}} |q_{l,\boldsymbol{i}}(x)|^2\right)^{1/2}}, & \text{otherwise} \end{cases}$$

$$(2.16)$$

and for the low frequency coefficients  $\tilde{q}_0(x)$  is given by

$$\tilde{q}_0(x) = \begin{cases} q_0(x), & \text{if } |q_0(x)| \le g(x) \\ g(x), & \text{otherwise} \end{cases}$$
(2.17)

We next focus on how to compute the resolvent  $(I + \partial G)^{-1}$ . Recall that G is defined as

$$G(\boldsymbol{u}) = \langle \boldsymbol{u}, \boldsymbol{f} \rangle + \delta_{\boldsymbol{S}}(\boldsymbol{u})$$

For a given  $\boldsymbol{v} = (v_1, \cdots, v_k)$ , we compute the resolvent  $\tilde{\boldsymbol{v}} = (I + \partial G)^{-1}(\boldsymbol{v})$  by

$$\begin{split} \tilde{\boldsymbol{v}} &= (I + \tau \partial G)^{-1}(\boldsymbol{v}) \\ &= \arg \min_{\boldsymbol{u}} \frac{\|\boldsymbol{u} - \boldsymbol{v}\|^2}{2\tau} + \langle \boldsymbol{u}, \ \boldsymbol{f} \rangle + \delta_{\boldsymbol{S}}(\boldsymbol{u}) \\ &= \arg \min_{\boldsymbol{u} \in \boldsymbol{S}} \|\boldsymbol{u} - (\boldsymbol{v} - \tau \boldsymbol{f})\|^2 \\ &= \prod_{\boldsymbol{S}} (\boldsymbol{v} - \tau \boldsymbol{f}) \end{split}$$

where  $\prod$  denotes the orthogonal projection onto the simplex constraint S. In the literature, there are some classic algorithms to compute the orthogonal projection onto a simplex constraint in  $\mathbb{R}^K$ , see [37]. For completeness, we describe the algorithm presented in a recent report [19] in Algorithm 2.

Algorithm	<b>2</b>	Pro	iection	onto	a	simplex	Δ	in	$\mathbb{R}^{I}$	ſ
-----------	----------	-----	---------	------	---	---------	---	----	------------------	---

**Require:** Input  $\vec{z} = (z_1, z_2, \dots, z_K) \in \mathbb{R}^K$ 

- 1: Sort  $z_i$  in ascending order,  $z_1 \leq z_2 \leq \ldots z_K$ , set i = K 1. 2: Compute  $t_i = \frac{\sum_{j=i+1}^{K} z_j 1}{K i}$ . If  $t_i > z_i$  set  $t^* = t_i$ , and go to 4. Otherwise, set  $i \leftarrow i 1$  and repeat 2. if i = 0, go to 3.
- 3: Set  $t^* = \frac{\sum_{j=1}^n z_j 1}{\sum_{j=1}^n z_j 1}$
- 4: Return  $x^* = (z t^*)^+$  as the projection.

To complete Algorithm 1, we also need to update

$$\bar{u}^{n+1} = 2u^{n+1} - u^n. \tag{2.18}$$

We have built in our model a collection of  $c_i$ , which represents the average intensity of each region. In doing so, we are actually using piecewise constant functions to approximate the original image. The framelet based segmentation model itself does not give any information about these constants, so we need some rough estimates to start with. We use the K-means algorithm to get these rough estimates. We note that although the mean value  $c_i$  is assumed given in the proposed model (2.8), in practice, the vector  $c_i$  can also be updated for a better estimation. According to the Mumford-Shah model (1.1), for a fixed segmentation  $\Omega_i$ ,

$$c_i = \frac{\sum_{x \in \Omega} I(x) u_k(x)}{\sum_{x \in \Omega} u_i(x)}.$$
(2.19)

Finally, suppose that we successfully solved the relaxed minimization problem (2.8)and that the solution  $\boldsymbol{u}^* = (u_1^*, \cdots, u_k^*)$  happens to be binary. It is then a solution to the original nonconvex problem (2.6). If it is not binary, we need to force the solution to be binary as in (1.7).

Overall, the complete wavelet frame based image segmentation model is present as Algorithm 3.

**2.4.** Analysis. It is important to analysis the relation between the convexified model (2.8) and the original nonconvex problem (2.6). As we declared previously, if the global minimum  $u^*$  of (2.8) is binary, then it is naturally a global minimum of the nonconvex model. However, the existence of a global binary solution is unknown. In [3], an analysis of binary solution under specific conditions for total variation based segmentation model (1.8) are provided. Here, we provide a similar analysis from the point of view of saddle point formulation.

We observe that the objective function  $E^{P}(\boldsymbol{u})$  in (2.8) is convex and the constraint set S is compact; thus the set of minima of  $E^{P}(u)$  is nonempty and compact by classical convex analysis. On the other hand, if we consider the dual model of the Algorithm 3 Multi-phase wavelet frame segmentation method

- **Require:** Initialization. Use the K-mean algorithm to obtain an initial guess of the mean intensity value for each region  $c_k$  for  $k = 1, \dots, K$ . Compute  $\boldsymbol{f}$  by (2.4). Choose  $\sigma > 0$  and  $\tau > 0$ .
- 1: for  $n \ge 0$ , update  $p^{n+1}, u^{n+1}, \bar{u}^{n+1}$  by do
- 2: Compute  $\mathbf{p}^{n+1} = \Pi_{\mathbf{Y}}(\mathbf{p}^n + \sigma \mathbf{W} \bar{\mathbf{u}}^n)$  where  $\Pi_{\mathbf{Y}}$  is given by (2.16) for high frequency coefficients and (2.17) for low frequency ones.
- 3: Compute  $\boldsymbol{u}^{n+1} = \prod_{\boldsymbol{S}} (\boldsymbol{u}^n \tau \boldsymbol{W}^\top \boldsymbol{p}^{n+1} \tau \boldsymbol{f})$  by computing the projection onto simplex set for each pixel using Algorithm 2.
- 4: Update  $\bar{\boldsymbol{u}}^{n+1} = 2\boldsymbol{u}^{n+1} \boldsymbol{u}^n$
- 5: Update the mean intensity value  $c_i$  by (2.19) for  $k = 1, \dots, K$ .
- 6: end for.
- 7: Obtain binary solution by (1.7)

primal dual form (2.8), we have

$$\max_{\boldsymbol{p}\in\boldsymbol{P}} \left\{ -F^{*}(\boldsymbol{p}) + \left( \min_{\boldsymbol{u}\in\boldsymbol{U}} \langle \boldsymbol{W}^{\top}\boldsymbol{p} + \boldsymbol{f}, \ \boldsymbol{u} \rangle + \delta_{\boldsymbol{S}}(\boldsymbol{u}) \right) \right\}$$
$$= \max_{\boldsymbol{p}^{(k)}\in\boldsymbol{Y}} \left\{ \min_{\boldsymbol{u}\in\boldsymbol{S}} \sum_{k=1}^{K} \langle \boldsymbol{W}^{\top}\boldsymbol{p}^{(k)} + f_{k}, \ \boldsymbol{u}_{k} \rangle \right\}$$
(2.20)

Denote

$$E(\boldsymbol{u}, \boldsymbol{p}) := \sum_{k=1}^{K} \langle \boldsymbol{W}^{\top} \boldsymbol{p}^{(k)} + f_k, \ u_k \rangle$$
(2.21)

$$E^{D}(\boldsymbol{p}) := \min_{\boldsymbol{u} \in \boldsymbol{S}} \sum_{k=1}^{K} \langle \boldsymbol{W}^{\top} \boldsymbol{p}^{(k)} + f_{k}, \ u_{k} \rangle$$
(2.22)

By classical convex analysis theory, the minimization problem (2.8) satisfies Slater's constraint qualification, and strong duality holds. In other words, we can exchange the order of the max and min in (2.12) and obtain the equivalent saddle point form

$$\max_{p^{(k)} \in \boldsymbol{Y}} \min_{\boldsymbol{u} \in \boldsymbol{S}} E(\boldsymbol{u}, \boldsymbol{p})$$
(2.23)

On the other hand, we may derive the close form of the dual function  $E^{D}(\mathbf{p})$ . For any  $\mathbf{v} = (v_1, \dots, v_K) \in \mathbf{S}$ , denote  $\vec{v}(x) = (v_1(x), \dots, v_K(x))$  as the vector at each pixel x, then  $\vec{v}(x)$  is in K-dimensional simplex constraint

$$\Delta_K := \{ \vec{z} = (z_1, \cdots, z_K) | \sum_{k=1}^K z_k = 1, z_k \ge 0 \}.$$

Therefore, for any vector  $\vec{z} = (z_1, \cdots, z_K) \in \mathbb{R}^K$  and  $\vec{w} \in \mathbb{R}^K$ , we have

$$\min_{\vec{z}\in\Delta_K}\sum_{k=1}^K v_i w_i = \min\left\{w_1,\cdots,w_K\right\}.$$

Thus, the above dual model (2.20) is

$$E^{D}(\boldsymbol{p}) = \left\langle \min_{k} \left\{ \boldsymbol{W}^{\top} \boldsymbol{p}^{(k)} + f_{k} \right\}, \ \boldsymbol{1} \right\rangle$$
(2.24)

where **1** is the all-one vector in  $\mathbb{R}^n$ .

Furthermore, there exist a primal solution  $u^*$  and a dual solution  $p^*$  that form a saddle point pair of (2.12), i.e.

$$E(\boldsymbol{u}^*, \boldsymbol{p}) \leq E(\boldsymbol{u}^*, \boldsymbol{p}^*) \leq E(\boldsymbol{u}, \boldsymbol{p}^*), \quad \forall \boldsymbol{u} \in \boldsymbol{S}, \quad \forall \boldsymbol{p}^{(k)} \in \boldsymbol{Y}$$

and

$$E^{P}(u^{*}) = E(u^{*}, p^{*}) = E^{D}(p^{*})$$

Conversely, if  $(\boldsymbol{u}^*, \boldsymbol{p}^*)$  is a saddle point of  $E(\boldsymbol{u}, \boldsymbol{p})$ , then  $(\boldsymbol{u}^*, \boldsymbol{p}^*)$  are solutions for the primal and dual problems respectively.

REMARK 1. Let  $(\boldsymbol{u}^*, \boldsymbol{p}^*)$ ,  $\boldsymbol{u}^* = (u_1^*, \cdots, u_K^*)$ ,  $\boldsymbol{p}^* = (p^{(1),*}, \cdots, p^{(K),*})$  be a saddle point pair of (2.23). If at some pixel location  $x \in \Omega$ , there exists a unique index  $k_0$ such that

$$k_0 = \arg\min_k \left\{ (\boldsymbol{W}^\top p^{(k),*})(x) + f_k(x) \right\}$$
(2.25)

then  $u^*(x)$  satisfies

$$u_k^*(x) = \begin{cases} 1 & \text{if } k = k_0 \\ 0, & \text{otherwise.} \end{cases}$$

If there are several  $k_1(x), \dots, k_J(x)$  that achieve the minimum condition, then

$$\sum_{j=1}^{J} u_{k_j}^*(x) = 1 \text{ and } u_k^*(x) = 0 \text{ for } k \notin \{k_1, \cdots, k_J\}$$

When  $p^*$  is obtained, we can see from (2.20) that minimal of  $E(u, p^*)$  over  $u \in S$  is achieved only when  $u_k(x) = 1$  for  $k = k_0$  if  $k_0$  is the only index achieving the maximum of the . Thus u is binary at  $x \in \Omega$ . Similar argument is applied for several minimum index.

REMARK 2. Note that the low frequency  $||W_0u_k||_1$  is also penalized in our model, which is not usual for wavelet frame based image restoration model. In fact, since  $u_k(x) \ge 0$  for all x and  $W_0(x)$  is a low pass average filter,  $(W_0u_k)(x) \ge 0$  for all x. We have

$$\sum_{k=1}^{K} \|g \cdot W_0 u_k\|_1 = \sum_{x \in \Omega} g(x) W_0 \sum_{k=1}^{K} u_k(x) = \|W_0 g\|_1$$

Since it is constant for a given g(x), the regularization  $\min \sum_{k=1}^{K} \|\mathbf{W}u_k\|_{1,2}$  is equivalent to  $\min \sum_{k=1}^{K} \|\mathbf{W}u_k\|_{1,2}^0$ . The proposed model is thus equivalent to the usual high frequency  $\ell_{1,2}$  penalization model. However, the penalization of low frequencies allows the algorithm to converge faster and also more likely yields binary solution than the one without low frequency during iterative steps. This will be further illustrated and explained in Section 3.6.

**3.** Numerical results. In this section, we give some numerical tests of the proposed model (2.8) and compare them with some other existing ones.

**3.1. Example 1: 2D piecewise constant images.** We first show some 2D toy examples. Given the input image corrupted by some noise, we want to label the image pixels into K regions. The framelet based model almost perfectly reconstructs the ground truth. In this example, segmentation is equivalent to denoising in some sense. In Figure 3.1, we show the segmentation result with 4 regions for a piecewise constant image, and after plugging back in the mean value of each region, it is very close to the ground truth image.



Fig. 3.1: Segmentation of a piecewise constant test image

In the following, we compare the performance of framelet method with other existing methods. Figure 3.2 shows the segmentation results with different algorithms on the same test image as Figure 3.1 but with much heavier noise. Alpha expansion and alpha-beta swap [4] are often considered as two state-of-the-art graph based methods using anisotropic discrete total variation (TV) in the model (1.1). The other method we draw into comparison is the smoothed dual total variation model proposed in [3], where isotropic total variation and a smoothed dual algorithm is applied. Note that in Figure 3.2, as well as Figure 3.3, 3.4, we do not implement these algorithms by ourselves but quote these results directly from [3] for a fair comparison.



Fig. 3.2: From left to right: (a) input image, (b) alpha expansion [4], (c) alpha-beta swap [4], (d) smoothed dual total variation model [3], (e) wavelet frame model

In Figure 3.3, we show another toy example with three circles. The image is segmented into 4 regions with different methods. In Table 3.1, we compare the percentage of misclassified pixels for the tests in Figure 3.2 and 3.3. Note that the results for other experiments are draw directly from Table 1 in [3]. This table shows that our

proposed method has lower misclassified rate for these two toy examples.



Fig. 3.3: From left to right: (a) input image, (b) ground truth, (c) alpha-expansion [4], (d) smoothed dual total variation [3], (e) wavelet frame model

Table 3.1 Rate of misclassified pixels

Test	Alpha expansion	Alpha-beta swap	Dual model	Framelet model
Figure 3.2	8.89%	6.12%	5.51%	4.93%
Figure 3.3	1.17%	1.17%	1.06%	1.03%

In Figure 3.4, we compare the segmentation results for a color image in 10 regions. For the three methods in comparison, the images with alpha-expansion [4] and smoothed dual total variation model are directly taken from [3]. The segmentation obtained by our method shows more details for the cloud in the sky and sharper edges for the leaves and flowers since our method can adapt to different scales present in the images.

**3.2. Example 2. Triple junction test.** The input image (a) in Figure 3.5 arises in the minimal partition test and it is often used to illustrate if the global solution can be found. The color of the pixels in the gray area of input image are unknown and to be filled in. By setting the data term  $f_i = 0$  for i = 1, 2, 3 inside the gray disk, and by the color distance as in (2.4) outside the gray area. The exact global solution to the Mumford-Shah model is known as triple junction such that the boundary meets with 120 degree angles in the center. For this example, the algorithms produce different results. Alpha expansion algorithm is able to produce a binary solution, but is not correct. The dual model produces a nearly correct binary result. The wavelet frame based model is not able to produce a binary result before the final thresholding step. However, the final thresholding successfully produce the right result, which in some sense justifies the thresholding scheme proposed above.

**3.3. Example 3. 3D Kidney vascular system.** This example shows the segmentation result on 3D image. The input image is a stack of slices of the kidney vascular system. Figure 3.6 shows the segmentation result rendered from 3 different angles using our home made software.

We also compare our results with other models in figure 3.7. We compare with the Chan-Vese active contour method  $[18]^1$  and smoothed dual total variation model [3].

 $<sup>^1{\</sup>rm Code}$  is downloaded from http://www.mathworks.com/matlabcentral/fileexchange/34548-active-contour-without-edge



Fig. 3.4: Color image. From left to right: (a) input image, (b) alpha-expansion [4] , (c) smoothed dual total variation [3], (d) wavelet frame model



Fig. 3.5: Triple junction test. From left to right: (a) input image, (b) alpha expansion [4], (c) smoothed dual total variation model [3], (d) wavelet frame model



Fig. 3.6: 3D view of kidney vascular system from different angles

FRAMELET BASED IMAGE SEGMENTATION



Fig. 3.7: Comparison of different models for two views. From left to right: Chan-Vese model [18], smoothed dual total variation model [3], the framelet model. From up to down: different 3D views.

**3.4. Example 4. Leaf vein system.** The input image is a patch taken from a simulated leaf vein system which involves some preprocessing. Here, we compare with the smoothed dual total variation method and a fuzzy level set method for medical image segmentation [34] (code is downloaded from the authors' webpage). These models all require in advance some constants  $c_i$  which roughly represents the average intensity of each region. In order to compare their performance in a fair manner, we set  $c_i$  equal for each method. Framelet based model extracts many more details than other compared models.

We also test the segmentation algorithm for a real leaf image, see Figure 3.10. Some preprocessing is also made to the patch, such as fixing lighting inhomogeneity. The constants  $c_i$  is also set to be the same for each method. Again, the framelet based model outperforms the other models.

**3.5. Efficiency and GPU acceleration.** Finally, we given some comments on the efficiency and convergence of the algorithm here. The major computation complexity lies in the framelet transform. Since the transform is redundant, the efficiency is lower compared with TV based models. For 2D images, each iteration steps is 4 to 6 times slower than the TV based models. In practice, wavelet frame based model usually takes very few steps to converge. In all the examples above, the algorithm takes less than 50 steps to meet the stopping criteria, (relative error less than  $10^{-4}$ ). TV based models usually take hundreds of steps to converge. In total, the time complexity is thus comparable to TV based models. The spatial storage,



Fig. 3.8: Simulated and real leaf vein system examples



Fig. 3.9: Simulated leaf vein system: (a)smoothed dual total variation model [3], (b)fuzzy level set model [34], (c) Chan-Vese model [18], (d) wavelet frame model.

however, is still several times higher than TV based models. To further speed up the algorithm, we use a GPU implementation. To compare the performance of GPU and CPU implementations, we use a test image of size 256\*256\*100 and do a two-phase segmentation. The time cost per step is 480.3ms for GPU (Nvidia Quadro 5000) and 16103ms for CPU(Xeon E5500\*2), the efficiency is improved by a factor of 33, see Figure 3.11.

**3.6.** Discussions on binary solutions. In this subsection, we give some intuitive explanation and numerical evidence why utilizing of low frequency has some advantages although the model is not changes as stated in Remark 2.

Denote the solution to a relaxed segmentation model by  $\boldsymbol{u}^* = \{\boldsymbol{u}_1^*, \boldsymbol{u}_2^*, \dots, \boldsymbol{u}_K^*\},$ 



Fig. 3.10: Segmentation results with real leaf image: (a) smoothed dual total variation model [3], (b) fuzzy level set model [34], (c) Chan-Vese model [18], (d) wavelet frame model.



Fig. 3.11: GPU acceleration

usually we expect them to have two properties. One is that each  $u_k^*$  should have some regularity yet retain sharp edges which can be achieved by doing thresholding in the high frequencies; the other is to be close to binary solution. Since  $u^*$  satisfy the simplex constraint at each pixel, the second property implies that the support of  $u_k$  should overlap as little as possible. Together, these two properties require that  $\boldsymbol{u}_k$  has sparsity not only in the high frequencies but also in the low frequencies in the framelet domain. The usual practice of  $\ell_1$  minimization is essentially to pursue the sparsity asymptotically by doing thresholding. This is essentially why we introduce the low frequencies in the  $\ell_1$  minimization model and do thresholding to both high and low frequencies. During the iteration process,  $u_k$  would be sparser but it won't vanish since the simplex constraint also means that the union of the support of  $u_I$  is  $\Omega$ . As a consequence, doing thresholding on the low frequencies only contributes to enlarging the difference among  $W^T p^n + u$  at every pixel  $x \in \Omega$ . An extreme point solution of  $u^n$  will be more likely to be obtained when projected onto the simplex. This is exactly the desired property that we want. For these reasons, we observe that in many cases, the solution is almost binary even before the final thresholding step. Figure 3.12 shows segmentation result before the final binary step (1.7). We can see that the solution is almost binary already.

Furthermore, the usage of low frequency  $\ell^1$  norm penalization can significantly speed up the convergence as shown in Figure 3.13. The acceleration is especially eminent at the beginning of the iteration. Once a binary solution is attained for some subregion, we observe that it usually tends to stay binary for the subsequent iterations, which accelerates the convergence.



Fig. 3.12: (a)original kidney image, (b)wavelet frame image segmentation solution  $u_1$ . The solution is almost binary before the binaryization step (1.7).

4. Conclusion. In this paper, we propose a multi-phase segmentation model based on wavelet frame. A primal-dual framework combined with a fast projection algorithm is applied to solve the proposed optimization problem. Numerical results show that the proposed model outperforms existing total variation and level set based models in segmentation of tubular structures, especially when the input image has different scales of structures. We also point out that unlike in applications such as image restoration and inpainting, in image segmentation problems, incorporating the



Fig. 3.13: Comparison of convergence rate

low frequencies of the framelet transform of an image has additional advantages, such as speed up the convergence of the algorithm and more likely yield a binary solution. However, theoretical analysis and connections between the global solution and the binary ones remain open and we will investigate along this direction in a future work.

Acknowledgments. The authors would like to thank Professor Dan Hu from INS, Shanghai Jiao Tong university for providing the images of simulated and real leaves. The authors would also like to thanks Professor Raymond Chan from Chinese University of Hong Kong for providing the data of kidney vascular system.

## REFERENCES

- S. ARIVAZHAGAN AND L. GANESAN, Texture segmentation using wavelet transform, Pattern Recognition Letters, 24 (2003), pp. 3197–3203.
- [2] E. BAE AND X.C. TAI, Efficient global minimization methods for image segmentation models with four regions. UCLA CAM report 11-82.
- [3] E. BAE, J. YUAN, AND X.C. TAI, Global minimization for continuous multiphase partitioning problems using a dual approach, International journal of computer vision, 92 (2011), pp. 112–129.
- [4] Y. BOYKOV AND V. KOLMOGOROV, An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision, Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.26, no.9, pp.1124-1137, Sept. 2004
- [5] X. BRESSON, S. ESEDOGLU, P. VANDERGHEYNST, J.P. THIRAN, AND S. OSHER, Fast global minimization of the active contour/snake model, Journal of Mathematical Imaging and Vision, 28 (2007), pp. 151–167.
- [6] E.S. BROWN, T.F. CHAN, AND X. BRESSON, Completely convex formulation of the chan-vese image segmentation model, International Journal of Computer Vision, (2011), pp. 1–19.
- [7] J.-F. CAI, B. DONG, S. OSHER AND Z. SHEN, Image restoration: total variation; wavelet frames; and beyond, Journal of the American Mathematical Society, 25 (4) (2012), 1033-1089.
- [8] J.F. CAI, S. OSHER, AND Z. SHEN, Split bregman methods and frame based image restoration, Multiscale modeling & simulation, 8 (2009), p. 337.
- X. CAI, R. CHAN, S. MORIGI, AND F. SGALLARI, Framelet-based algorithm for segmentation of tubular structures, Scale Space and Variational Methods in Computer Vision, (2012), pp. 411–422.

- [10] E.J. CANDES AND D.L. DONOHO, New tight frames of curvelets and optimal representations of objects with piecewise c2 singularities, Communications on pure and applied mathematics, 57 (2004), pp. 219–266.
- [11] A. CHAI AND Z. SHEN, Deconvolution: A wavelet frame approach, Numerische Mathematik, 106 (2007), pp. 529–587.
- [12] A. CHAMBOLLE AND T. POCK, A first-order primal-dual algorithm for convex problems with applications to imaging, Journal of Mathematical Imaging and Vision, 40 (2011), pp. 120– 145.
- [13] A. CHAMBOLLE AND D. CREMERS AND T. POCK, A convex approach to minimal partitions, Technical Report, Ecole Polytechnique, France (2011)
- [14] R.H. CHAN, T.F. CHAN, L. SHEN, AND Z. SHEN, Wavelet algorithms for high-resolution image reconstruction, SIAM Journal on Scientific Computing, 24 (2003), pp. 1408–1432.
- [15] R.H. CHAN, S.D. RIEMENSCHNEIDER, L. SHEN, AND Z. SHEN, Tight frame: an efficient way for high-resolution image reconstruction, Applied and Computational Harmonic Analysis, 17 (2004), pp. 91–115.
- [16] T.F. CHAN, S. ESEDOGLU, AND M. NIKOLOVA, Algorithms for finding global minimizers of image segmentation and denoising models, SIAM Journal on Applied Mathematics, (2006), pp. 1632–1648.
- [17] T. CHAN AND L. VESE, An active contour model without edges, Scale-Space Theories in Computer Vision, (1999), pp. 141–151.
- [18] T.F. CHAN AND L.A. VESE, Active contours without edges, Image Processing, IEEE Transactions on, 10 (2001), pp. 266–277.
- [19] Y. CHEN AND X. YE, Projection onto a simplex, Arxiv preprint arXiv:1101.6081, (2011).
- [20] I. DAUBECHIES, Ten lectures on wavelets, vol. 61, Society for Industrial Mathematics, 1992.
- [21] I. DAUBECHIES, B. HAN, A. RON, AND Z. SHEN, Framelets: MRA-based constructions of wavelet frames, Applied and Computational Harmonic Analysis, 14 (2003), pp. 1–46.
- [22] B. DONG, A. CHIEN, AND Z. SHEN, Frame based segmentation for medical images, Communications in Mathematical Sciences, 9 (2011), pp. 551–559.
- [23] B. DONG AND Z. SHEN, MRA-based wavelet frames and applications: image segmentation and surface reconstruction, survey paper for SPIE Defense, Security, and Sensing, 2012.
- [24] B. DONG AND Z. SHEN, MRA based wavelet frames and applications, IAS Lecture Notes Series, Summer Program on The Mathematics of Image Processing, Park City Mathematics Institute, (2010).
- [25] E. ESSER, X. ZHANG, AND T. F. CHAN, A general framework for a class of first order primaldual algorithms for convex optimization in imaging science, SIAM J. Imaging Sci., 3 (2010), pp. 1015–1046.
- [26] M.A.T. FIGUEIREDO AND R.D. NOWAK, An EM algorithm for wavelet-based image restoration, Image Processing, IEEE Transactions on, 12 (2003), pp. 906–916.
- [27] —, A bound optimization approach to wavelet-based image deconvolution, in Image Processing, 2005. ICIP 2005. IEEE International Conference on, vol. 2, IEEE, 2005, pp. II–782.
- [28] E. FRANCHINI, S. MORIGI, AND F. SGALLARI, Composed segmentation of tubular structures by an anisotropic PDE model, Scale Space and Variational Methods in Computer Vision, (2009), pp. 75–86.
- [29] E. FRANCHINI, S. MORIGI AND F. SGALLARI, Segmentation of 3D tubular structures by a PDEbased anisotropic diffusion model, Mathematical Methods for Curves and Surfaces, (2010), pp. 224–241.
- [30] T. GOLDSTEIN AND S. OSHER, The split bregman method for l1 regularized problems, SIAM Journal on Imaging Sciences, 2 (2009), pp. 323–343.
- [31] M. KASS, A. WITKIN, AND D. TERZOPOULOS, Snakes: Active contour models, International journal of computer vision, 1 (1988), pp. 321–331.
- [32] R. KIMMEL, Fast edge integration, Geometric Level Set Methods in Imaging, Vision, and Graphics, (2003), pp. 59–77.
- [33] R. KIMMEL AND A.M. BRUCKSTEIN, Regularized laplacian zero crossings as optimal edge integrators, International Journal of Computer Vision, 53 (2003), pp. 225–243.
- [34] B.N. LI, C.K. CHUI, S. CHANG, AND SH ONG, Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation, Computers in Biology and Medicine, 41 (2011), pp. 1–10.
- [35] R. MALLADI, J.A. SETHIAN, AND B.C. VEMURI, Shape modeling with front propagation: A level set approach, Pattern Analysis and Machine Intelligence, IEEE Transactions on, 17 (1995), pp. 158–175.
- [36] T. MCINERNEY AND D. TERZOPOULOS, Deformable models in medical image analysis, in Mathematical Methods in Biomedical Image Analysis, 1996., Proceedings of the Workshop on,

IEEE, 1996, pp. 171–180.

- [37] C. MICHELOT, A finite algorithm for finding the projection of a point onto the canonical simplex of n, Journal of Optimization Theory and Applications, 50 (1986), pp. 195–200.
- [38] D. MUMFORD AND J. SHAH, Optimal approximations by piecewise smooth functions and associated variational problems, Communications on pure and applied mathematics, 42 (1989), pp. 577–685.
- [39] J. J. MOREAU, Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. France, 93, 1965, pp. 273-299.
- [40] A. RON AND Z. SHEN, Affine system in L<sub>2</sub>(R<sup>d</sup>) I: the analysis of the analysis operator, J. Func. Anal., 148 (1997), pp. 408-447.
- [41] A. RON AND Z. SHEN, Affine systems in L<sub>2</sub>(ℝ<sup>d</sup>) II: Dual systems, Journal of Fourier Analysis and Applications, 3 (1997), pp. 617–637.
- [42] L. I. RUDIN, S. OSHER AND E. FATEMI, Nonlinear total variation based noise removal algorithms, Physica D. (1992) 60 259-68.
- [43] B. SANDBERG AND T.F. CHAN, A logic framework for active contours on multi-channel images, Journal of Visual Communication and Image Representation, 16 (2005), pp. 333–358.
- [44] Z. SHEN, Wavelet frames and image restorations, Proceedings of the International congress of Mathematicians, Vol IV, Hyderabad, India, (2010), Hindustan book Agency, (Rajendra Bhatia eds), 2834-2863.
- [45] J.L. STARCK, M. ELAD, AND D.L. DONOHO, Image decomposition via the combination of sparse representations and a variational approach, Image Processing, IEEE Transactions on, 14 (2005), pp. 1570–1582.
- [46] M. UNSER, Texture classification and segmentation using wavelet frames, Image Processing, IEEE Transactions on, 4 (1995), pp. 1549–1560.
- [47] W. YIN, S. OSHER, D. GOLDFARB, AND J. DARBON, Bregman iterative algorithms for l1minimization with applications to compressed sensing, SIAM Journal on Imaging Sciences, 1 (2008), pp. 143–168.
- [48] C. ZACH, D. GALLUP, J.-M. FRAHM AND M. NIETHAMMER, Fast global labeling for realtime stereo using multiple plane sweeps, in Vision, Modeling and Visualization Workshop (VMV), IOS Press, 2008.
- [49] X. ZHANG, M. BURGER AND S. OSHER, A unified primal-dual algorithm framework based on Bregman iteration, J. Sci. Comput. (2011) 46 20-46.
- [50] M. ZHU AND T. CHAN, An efficient primal-dual hybrid gradient algorithm for total variation image restoration, UCLA CAM Report, (2008), pp. 08–34.