Frame Based Segmentation for Medical Images

Bin Dong\textsuperscript{a}, Aichi Chien\textsuperscript{b}, Zuowei Shen\textsuperscript{c}

\textsuperscript{a}Department of Mathematics, University of California, San Diego, San Diego, 9500 Gilman Drive, La Jolla, CA, 92093-0112
\textsuperscript{b}Division of Interventional Neuroradiology, David Geffen School of Medicine at UCLA, 10833 LeConte Ave, Los Angeles, CA, USA
\textsuperscript{c}Department of Mathematics, National University of Singapore, Science Drive 2, Singapore, 117543.

Abstract

Medical image segmentation is an important but difficult problem that attracts tremendous attentions of researchers from various fields. In this paper, we propose a frame based model, as well as a fast implementation, for general medical image segmentation problems. Our model combines ideas of the frame based image restoration model of [1] with ideas of the total variation based segmentation model of [2, 3]. Numerical experiments show that the proposed frame based model outperforms total variation based model in terms of capturing key features of biological structures. Successful segmentations of blood vessels and aneurysms in 3D CT angiography images are also presented.

Keywords: Image segmentation, level set method, sparse approximation, tight frames.

1. Introduction

Segmenting biological structures, e.g. cortical or subcortical structures, blood vessels, tumors etc., from various types of medical images is very important for detecting abnormalities, studying and tracking progress of diseases,

\textsuperscript{1}Aichi Chien is supported in part by NSF CCF-0830554.
\textsuperscript{2}Zuowei Shen is supported in part by Grant R-146-000-113-112 from the National University of Singapore.

\textsuperscript{Email addresses: dongbin725@gmail.com (Bin Dong), aichi@ucla.edu (Aichi Chien), matzuows@nus.edu.sg (Zuowei Shen)}

Preprint submitted to . May 18, 2010
and surgery planning. Medical image segmentation is a difficult problem due to the fact that medical images commonly have poor contrasts, different types of noise, and missing or diffuse boundaries. There are numerous algorithms developed in the literature targeting either general segmentation problems or the segmentation of specific biological structures (see [4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and the references therein). In this paper, we propose a novel segmentation model that is based on tight frames and the fact that they provide a sparse approximation to piecewise smooth functions like images.

The theory of frames, especially tight frames and framelets (wavelet tight frames), were extensively studied in the past twenty years (see e.g. [14, 15, 16, 17, 18]), including the conditions on which a system of functions forms a tight frame, and the constructions of framelets. Examples of tight frames are translation invariant wavelets [19], curvelets [20], and framelets [15]. In contrast to orthogonal bases, tight frames give redundant representations to signals and images. The redundancy of tight frames usually leads to sparse approximation of images, which is known to be a desirable property for image restoration problems, like denoising, inpainting, deblurring, etc. [19, 20, 1, 21, 22, 23, 24, 25]. There are also some research on texture classification and segmentation using wavelets or wavelet frames [26, 27]. However, to our best knowledge, utilizing the property of sparse approximation of tight frames for image segmentation problems has not been considered in the literature.

The rest of the paper is organized as follows. In Section 2, we will first briefly review the concept of frames and framelets, and then introduce our frame based segmentation model together with its fast implementation. Numerical comparisons and results will be given in Section 3, and concluding remarks will be given in the last section.

2. Frame Based Segmentation Model

2.1. Frames and Framelets

In this subsection, we will briefly introduce the concept of tight frames and framelets. Interesting readers should consult [14, 15, 16] for theories of frames and framelets, [17] for a short survey on theory and applications of frames, and [18] for a more detailed survey.

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

\[
f = \sum_{h \in X} (f, h) h \quad \forall f \in L_2(\mathbb{R}),
\]  

(2.1)
where \(\langle \cdot, \cdot \rangle\) is the inner product of \(L^2(\mathbb{R})\). For given \(\Psi := \{\psi_1, \ldots, \psi_r\} \subset L^2(\mathbb{R})\), the affine (or wavelet) system is defined by the collection of the dilations and the shifts of \(\Psi\) as

\[
X(\Psi) := \{\psi_{\ell,j,k} : 1 \leq \ell \leq r; \ j, k \in \mathbb{Z}\} \quad \text{with} \quad \psi_{\ell,j,k} := 2^{j/2} \psi_\ell(2^j \cdot -k).
\]

(2.2)

When \(X(\Psi)\) forms a tight frame of \(L^2(\mathbb{R})\), it is called a tight wavelet frame, and \(\psi_\ell, \ell = 1, \ldots, r\), are called the (tight) framelets.

To construct a set of framelets, usually, one starts from a compactly supported refinable function \(\phi \in L^2(\mathbb{R})\) (a scaling function) with a refinement mask \(h_0\) satisfying

\[
\hat{\phi}(2 \cdot) = \hat{h}_0 \hat{\phi}.
\]

Here \(\hat{\phi}\) is the Fourier transform of \(\phi\), and \(\hat{h}_0\) is the Fourier series of \(h_0\) with \(\hat{h}_0(0) = 1\), which means that a refinement mask of a refinable function must be a lowpass filter. For a given compactly supported refinable function, the construction of a tight framelet system is to find a finite set \(\Psi\) that can be represented in the Fourier domain as

\[
\hat{\psi}_\ell(2 \cdot) = \hat{h}_\ell \hat{\phi}
\]

for some 2\(\pi\)-periodic \(\hat{h}_\ell\). The unitary extension principle (UEP) of [15] says that \(X(\Psi)\) in (2.2) generated by \(\Psi\) forms a tight frame in \(L^2(\mathbb{R})\) provided that the masks \(h_\ell\) for \(\ell = 0, 1, \ldots, r\) satisfy

\[
\sum_{\ell=0}^r \hat{h}_\ell(\xi) \overline{\hat{h}_\ell(\xi + \gamma \pi)} = \delta_{\gamma,0}, \quad \gamma = 0, 1,
\]

(2.3)

for almost all \(\xi\) in \(\mathbb{R}\). While \(h_0\) corresponds to a lowpass filter, \(\{h_\ell; \ell = 1, 2, \ldots, r\}\) must correspond to highpass filters by the UEP. The sequences of Fourier coefficients of \(\{h_\ell; \ell = 1, 2, \ldots, r\}\) are called framelet masks. In our implementation, we adopt the piecewise linear B-spline framelet constructed in [15]. The refinement mask is \(h_0(\xi) = \cos^2(\xi/2)\), whose corresponding lowpass filter is \(h_0 = \frac{1}{4}[1, 2, 1]\). Two framelets are \(\hat{h}_1 = -\frac{\sqrt{2}}{2} \sin(\xi)\) and \(\hat{h}_2 = \sin^2(\xi/2)\), whose corresponding highpass filters are

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

The associated refinable function and framelets are given in Figure 1.
Figure 1: Piecewise linear refinable spline and framelets.

With a one-dimensional framelet system for \( L_2(\mathbb{R}) \), the \( s \)-dimensional framelet system for \( L_2(\mathbb{R}^s) \) can be easily constructed by tensor products of one-dimensional framelets (see e.g. [14, 18]). If we have one scaling function and \( r \) tight framelets in 1D, then after tensor product, we obtain a tight frame system generated by one scaling function and \( r^s - 1 \) tight framelets. For a given function \( f \in L_2(\mathbb{R}^s) \), the \( L \)-level framelet decomposition of \( f \) is the combination of the coefficients \( \{ \langle f, 2^{-L/2}\phi(2^{-L} \cdot -j) \rangle \} \) at the prescribed coarsest level \( L \), and the framelet coefficients

\[
\{ \langle f, 2^{-l/2}\psi_i(2^{-l} \cdot -j) \rangle, 1 \leq i \leq r^s - 1 \}
\]

for \( 0 \leq l \leq L \).

In the discrete setting, a discrete image \( f \) is considered as the coefficients \( \{ f_i = \langle f, \phi(\cdot - i) \rangle \} \) up to a dilation, where \( \phi \) is the refinable function associated with the framelet system, and \( \langle \cdot, \cdot \rangle \) is the inner product in \( L^2(\mathbb{R}^2) \). The \( L \)-level discrete framelet decomposition of \( f \) is then the coefficients \( \{ \langle f, 2^{-L/2}\phi(2^{-L} \cdot -j) \rangle \} \) at a prescribed coarsest level \( L \), and the framelet coefficients are

\[
\{ \langle f, 2^{-l/2}\psi_i(2^{-l} \cdot -j) \rangle, 1 \leq i \leq r^2 - 1 \}
\]

for \( 0 \leq l \leq L \).

In the discrete setting, an \( s \)-dimensional image, which is an \( s \)-dimensional array, can be understood as a vector living in \( \mathbb{R}^n \), with \( n \) the total number of pixels in the image. For simplicity of notations, we represent the framelet decomposition and reconstruction as matrix multiplications \( Wu \) and \( W^Tv \) respectively. Here \( W \in \mathbb{R}^{k \times n} \) satisfies \( W^TW = I \), i.e. \( u = W^TWu, \forall u \in \mathbb{R}^n \), by the unitary extension principle [15]. In our numerical implementation, however, we use the fast tensor product tight wavelet frame decomposition and reconstruction algorithms of [18, 25].
We now introduce some additional notations that we will use throughout this paper. Let $W_0$ be the submatrix of $W$ that corresponds to the decomposition with respect to the refinable function; and denote $W_{l,i}$ with $1 \leq l \leq L$ and $1 \leq i \leq r^s - 1$, the submatrix of $W$ that corresponds to the decomposition at the $l$-th level with respect to the $i$-th framelet. Under this notation, $W$ can be written as

$$W = \begin{pmatrix} W_0 \\ (W_{l,i}) \end{pmatrix} = \begin{pmatrix} W_0 \\ W_{1,1} \\ W_{1,2} \\ \vdots \\ W_{L,r^s-1} \end{pmatrix}.$$ 

In this paper we shall use the tight framelet transforms without downsampling [25]. In this case, all $W_{l,i}$ and $W_0$ have the same number of rows, and we denote that number as $m$. Furthermore, all vectors in $\mathbb{R}^n$ are taken to be column vectors by convention. For any two vectors $v$ and $w$ in $\mathbb{R}^n$, we denote $v^\top w$ the inner product of $v$ and $w$, and denote $vw$ the component-wise multiplication of $v$ and $w$, i.e. $(vw)(j) = v(j)w(j), \forall j = 1, \ldots, n$.

2.2. Segmentation Model

Our frame based segmentation model (2.4) is motivated by the total variation based (TV-based) segmentation model of [2, 3]. Interesting readers should consult [2, 3] for more details.

For a given image $f \in \mathbb{R}^n$, we consider the following optimization problem

$$\min_{0 \leq u \leq 1, c_1, c_2} \|g_W \cdot W u\|_1 + \mu r(c_1, c_2)^\top u,$$

where $\| \cdot \|_1$ denotes the $\ell_1$-norm and $r(c_1, c_2) = (c_1 - f)^2 - (c_2 - f)^2$ where $c_1$ and $c_2$ are constant vectors.

Here $g_W$ is a diagonal matrix defined as

$$g_W = \text{diag}\{0^\top, v_{1,1}^\top, v_{1,2}^\top, \ldots, v_{1,r^s-1}^\top, \ldots, v_{L,r^s-1}^\top\},$$

where $v_{l,i} \in \mathbb{R}^m$ and $0 \in \mathbb{R}^m$. Then $g_W \cdot W u$ can be written as

$$g_W \cdot W u = \sum_{l,i} v_{l,i}(W_{l,i}u).$$
There are multiple ways to choose the weight function $v_{l,i}$. In this paper we choose $v_{l,i} = v$ for all $l$ and $i$, where

$$v(j) = \frac{1}{1 + \sigma \sum_{i=1}^{r-1} |(W_{l,i} f)(j)|^2} \quad \text{for } j = 1, 2, \ldots, m.$$ 

Notice that $g_W$ can be regarded as the edge indicator function under the framelet transform $W$.

To solve (2.4), one can alternatively optimize variables $u$ and $c_i$, $i = 1, 2$. Since when $u$ is fixed, the optimal values $c_i$ can be easily determined. Therefore, the key step is to optimize (2.4) with $c_i$ fixed. Here we adopt the idea of [28] by using split Bregman iteration [29] to solve the TV-based segmentation model [2, 3]. Following a similar derivation and using the fact that $W^T W = I$, we obtain the following algorithm for (2.4):

$$
\begin{align*}
    u_{k+\frac{1}{2}} &= W^T (d^k - b^k) - \frac{\mu}{\lambda} r(c_1^{k+1}, c_2^{k+1}) \\
    u^{k+1} &= \max\{\min\{u^{k+\frac{1}{2}}, 1\}, 0\} \\
    d^{k+1} &= T_{\gamma W/\lambda} (W u^{k+1} + b^k) \\
    b^{k+1} &= b^k + (W u^{k+1} - d^{k+1}) \\
    c_{1}^{k+1} &= M(f, \Omega^{k+1}), \quad c_{2}^{k+1} = M(f, (\Omega^{k+1})^c), \quad \Omega^{k+1} = \{u^{k+1} > \alpha\}.
\end{align*}
$$

In the above algorithm, $\mu$ is the parameter as in (2.4), $\lambda$ is another parameter that comes from Bregman iteration, $\alpha \in [0, 1]$, $T_{\delta}$ is the soft-thresholding operator defined as

$$
(T_{\delta}(x))(j) := \begin{cases} 
    x(j) - \delta(j), & \text{if } x(j) > \delta(j) \\
    0, & \text{if } -\delta(j) \leq x(j) \leq \delta(j) \\
    x(j) + \delta(j), & \text{if } x(j) < -\delta(j),
\end{cases}
$$

and $M(f, \Omega)$ returns the mean value of $f$ within domain $\Omega$. After we obtain a solution $u^*$ from the algorithm (2.5), the segmentation of image $f$ is given by the $\alpha$ level set of $u^*$. It is proven for TV-based model that any $\alpha$ level set of $u^*$, for almost all $\alpha \in [0, 1]$, gives a meaningful segmentation of image $f$ (see [2, 3] for details). Although we do not have a similar theory for the frame based model (2.4) yet, our numerical results support a similar conclusion. In our experiments, $\alpha$ is taken to be $0.5$.

Note that (2.5) is a very efficient algorithm. For each iteration $k$, the most time consuming operation is performing fast framelet decomposition.
and reconstruction as suggested by [16], i.e. \( W \) and \( W^\top \), which are of the same complexity as fast Fourier transform (FFT). Furthermore, our numerical experiments show that usually we only need a few hundred iterations until the algorithm converges. We leave the details to the next section.

**Remark 2.1.** The crucial difference between our model (2.4) from the TV-based model [2, 3] is that we are penalizing the \( \ell_1 \)-norm of the framelet coefficients \( W u \) instead of \( |\nabla u| \). The advantages of using tight frames over TV are as follows:

1. Functions that are sparse under operator \( \nabla \) are piecewise constant functions. Since in general a level set function (the variable \( u \) in (2.4)) can be any piecewise smooth function, and piecewise smooth functions have sparser representations under tight frame systems, penalizing \( \ell_1 \)-norm of \( W u \) generally should generates better results than penalizing \( |\nabla u| \) as confirmed by researches in image restoration problems [1, 20, 21, 22, 23, 24].

2. On the other hand, \( \|W u\|_1 \) contains more geometric information of the level sets of \( u \) than \( \|\nabla u\|_1 \). In particular for the case \( s = 2 \), when piecewise linear tight frames are used, we have not only a first order difference operator in the system, but also 2nd-4th order difference operators, that provides rich geometric information of the level sets of \( u \).

3. **Numerical Results**

In this section, we will first compare the frame based segmentation model (2.4) with the TV-based segmentation model [2, 3]. Then we show some segmentation results for 3D CT angiography (CTA) images using (2.4).

In our implementation, we adopt the stopping criteria: \( \|b^{k+1}\| < 10^{-3} \). Based on this stopping criteria, the number of iterations for 2D and 3D cases varies from 100 to 500. Within each iteration, the comparably expensive operation is the framelet decomposition and reconstruction, i.e. \( W \) and \( W^\top \). Although the complexity of applying \( W \) and \( W^\top \) is of the same order as FFT by applying the fast algorithm of [16], in practice the constant matters. We note, however, that this constant is not big, and hence the framelet decomposition and reconstruction can be done rather efficiently. For example, for a 3D image of size \( 50 \times 50 \times 50 \), the computational time for one level of framelet decomposition and reconstruction is approximately 5–6 times of the
computational time of forward and inverse FFT. This comparison is done using Matlab2007. Throughout this section, the parameter $\alpha$ in (2.5) is chosen to be 0.5, and the level of framelet decomposition $L$ is chosen to be 2.

3.1. Comparison of Frame Based Model with TV-based Model

We shall use the fast algorithm proposed by [28] to solve the TV-based model [2, 3]. The algorithm of [28] is based on the split Bregman algorithm [29] which is a rather efficient algorithm in solving $\ell_1$ regularization problems. Here we shall skip the details and refer the interesting readers to [28, 29].

Basically if one replace the update for $u^{k+\frac{1}{2}}$ in the first line of (2.5) by

$$\Delta u^{k+\frac{1}{2}} = \nabla \cdot (d^k - b^k) - \frac{\mu}{\lambda} r(c_1^k, c_2^k),$$

replace all $W$ by $\nabla$ and $g_W$ by $g$, then we obtain the algorithm proposed by [28]. Note that the Laplace equation above is solved by FFT, instead of Gauss-Seidel relaxation as proposed in [28].

In order to truly show the improvement of using tight frame systems, we pick the same set of parameters $(\mu, \lambda)$ and use the same $g_W$ for both models. As one can see from both Fig. 2 and Fig. 3 that using tight frame systems we can capture more features from the images and obtain better segmentations. Note that in the experiments, the ratio $\mu/\lambda$ is fixed due to the way that $u^{k+\frac{1}{2}}$ is updated.

![Figure 2: Comparisons of TV-based segmentation model (blue) with our frame based segmentation model (2.4) (red). The parameters $(\mu, \lambda)$ we used from column 1-3 are (20, 2), (30, 3) and (40, 4) respectively.](image-url)
3.2. 3D Segmentation Results

In this section we present the segmentation results for several 3D CTA images of brain blood vessels with aneurysms. Blood vessel segmentation and visualization is important for clinical tasks such as diagnosis of vascular diseases and blood flow simulation [30]. There are numerous methods developed for vessel segmentation in the literature (see e.g. [11, 12, 13] and the references therein). We note, however, that our model is entirely general and can be applied to any other type of medical images, e.g. MR images, with different biological structures.

Fig. 4 presents the segmented surfaces together with their corresponding axial, sagittal and coronal views. For all of the different subjects, we used the same set of parameters, i.e. \((\mu, \lambda) = (200, 10)\). The number of iterations for the four objects are 120, 163, 195 and 169 respectively. As one can see from Fig. 4 that crucial structures of the blood vessels and the aneurysms are well captured. It is also worth noticing that our model seems to be rather robust in terms of choice of parameters and changes of image contrasts.

4. Conclusion

In this paper we proposed a novel frame based segmentation model. Our numerical results showed the advantage of employing tight frame transforms
Figure 4: Row 1-4 shows segmentation results for the four different subjects. Column 2-4 present the axial, sagittal and coronal slices of the 3D image superimposed with the segmentation results.

in the energy functional over the traditional total variation. This is essentially because tight frames can provide sparser approximation to piecewise smooth functions and grant richer geometric information than total variation. Our results have shown that the frame based segmentation method can capture detailed structure of the vessels and aneurysms, and may be useful to assist medical evaluation.

References


