

# Isotropic Multiresolution Analysis for 3D-Textures and Applications in Cardiovascular Imaging

Simon K. Alexander, Robert Azencott, Manos Papadakis

Department of Mathematics, University of Houston, 651 Phillip G. Hoffman Hall, Houston,  
TX 77204-3008

## ABSTRACT

The main goal of this paper is to introduce formally the concept of texture segmentation/identification in three dimensional images. A major problem in texture texture segmentation/identification is the lack of robustness to both translations and rotations. This problem is more difficult to overcome in 3D-images, such as those generated by modalities such as x-ray CT and MRI. To facilitate 3D-texture segmentation/identification which is robust to 3D rigid motions we formally introduce the concept of steerable feature maps and of appropriate metrics in the feature space. We also introduce a new multiscale representation giving rise to a steerable feature map used in our exploratory project in cardiovascular imaging and we propose a 3D-texture segmentation algorithm utilizing this steerable feature map.

## 1 Introduction

In the recent years diagnostic techniques in medicine have been employing an increasing number of imaging modalities. This generates an enormous amount of data that radiologists and technicians have to process. Often critical diagnostic decisions must be made in areas where the quality of a medical image does not allow an unambiguous diagnosis. This is true despite the number of modalities and significant improvements in resolution and overall image quality.

Healthcare practice is increasingly emphasizing early diagnosis in hopes earlier treatment is more effective. This trend motivates the development of assistive diagnostic tools that will allow physicians to spot more easily regions of images indicating future risk to the patient's health. However, such regions are often not large enough to allow easy detection of may be even latent in an image cluttered by various types of noise that are difficult to model (and thus remove) in a post-acquisition stage.

Even when a region of interest is easily detectable, it's boundaries may not be well defined. To illustrate this concept we refer to the case of lipid concentrations inside arterial walls. We can locate such concentrations but the decision to treat each one of them depends primarily on the fibrous cap that seals such a lesion from the lumen. Once this cap is less than  $150\mu m$ , then an acute myocardial infarction may be imminent. This demonstrates the necessity for an accurate circumscription of the lipid pool from its sealing fibrous cap. This is, however, a difficult problem since the  $150\mu m$  length is very close to the resolution of intravascular ultrasound probes (IVUS) and much less than the resolution of other relevant modalities used in current clinical practice. In addition to this problem, removal of noise from the image is a risky approach since noise models are not clear for some modalities, e.g. ultrasound, and a prior knowledge of the true image is not available. For these reasons post-acquisition noise removal is not trusted by many medical practitioners. Overall the goal is to detect as accurately as possible various tissue types of diagnostic relevance. Particularly relevant methods to address this problem are texture segmentation/identification. Examples of data sets are given in Figure 1.

The biomedical images we are interested in are three-dimensional since they are generated by modalities such as x-ray CT, MRI and IVUS. Post-processing them in their original dimensionality increases the sensitivity of the texture segmentation/identification since (2D) slice by slice processing decorrelates the information in adjacent slices. The particular problem we are addressing with in this paper is 3D-texture segmentation. The identification problem is significantly more complicated and perhaps less attractive for a medical application, since the final diagnosis will come from a physician and not from a machine. *Segmenting different tissue types in their native dimensionality and circumscribing as accurately as possible their boundary surfaces assists diagnostic imaging especially when latent tissues are picked up by this process. This segmentation process must have the flexibility to account for the wide natural variation that may occur in tissues of the same type and for the often unknown variability due to noise, motion and other factors affecting the image acquisition.*

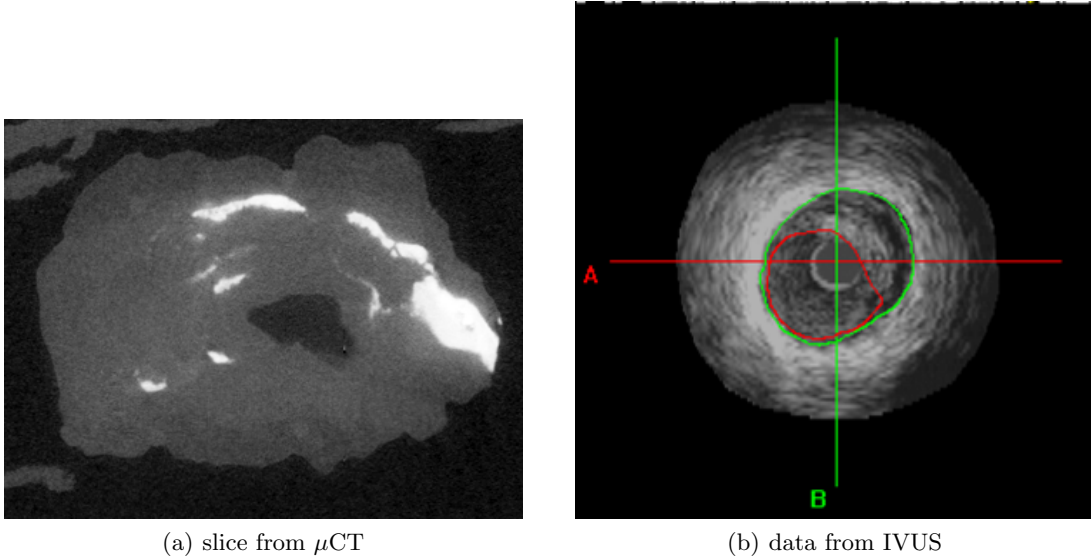


Figure 1: Examples of medical data.

## 2 3D Texture segmentation

In this work we are particularly interested in the segmentation of tissue types in three dimensional medical image data (MRI, x-ray CT, IVUS, etc). Texture segmentation is a large area with a rich literature, particularly in the 2D case. While some very general approaches to clustering and segmentation are known, state-of-the-art performance in applications is reached by tailoring each method to the particulars of the data representation and segmentation task. We will broadly constrain ourselves to discussing approaches that have been successfully applied in applications similar to our own — there are a very large number of segmentation algorithms in the literature, not all of which are applicable.

3D segmentation is significantly less developed than its 2D counterpart. Part of this is due to the reduced complexity of the 2D problem (which remains very difficult to do well), and also the computational demands of 3D processing. Perhaps another reason is that only more recently have 3D volumes become so important to clinical medical practice — demanding improvements in all aspects of handling and processing 3D data sets.

There have been several different approaches applied effectively to particular 3D segmentation problems. Without attempting any sort of complete categorization of the approaches, we will mention different broad classes of important algorithms.

Active contour/deformable models have become a popular way of asserting (typically physics based) models with constrained shape, e.g. smoothness constraints. *Snakes*<sup>1</sup> became very popular for both 2D and 3D applications, and a building block for many other approaches including hybrid approaches.<sup>2,3</sup>

Stochastic segmentation methods enjoyed broad popularity after the seminal paper of Geman and Geman introducing the idea of Markov Random Fields to the imaging community.<sup>4</sup> Since that time, a myriad of related Random Field approaches have been tried, typically through relaxation of a Bayesian model with various priors asserting things like boundary smoothness.<sup>5,6</sup> Much of this literature *particularly in 3D* is concerned with addressing the computational challenges of stochastic relaxation.

Another influential direction introduced in 1988 were approaches based on *isosurfaces*, the *level sets* method of Osher and Sethian.<sup>7</sup> Here isosurfaces serve as an alternative to parametrized models, and level-sets are (numerical) techniques to track evolution of interfaces of an implicit surface model with the measurements. Such methods have been further developed<sup>8,9</sup> and applied to three dimensional segmentation problems.

Other approaches include ones based on optical flow,<sup>10</sup> graph based approaches,<sup>11,12</sup> morphological methods,<sup>13,14</sup> and learning algorithms.

Each of these methods has strengths and challenges, particularly when applied to large, three dimensional data sets. For our applications of interest, we have identified the following issues as key:

- Robustness to changes in orientation of the data.
- Ability to discriminate subtly different tissues (textures) with random local orientation.
- Able to utilize high dynamic range medical data, even in the presence of significant noise.
- Computationally reasonable on volumetric data of at least  $512 \times 512 \times 512$  voxels.

There is another issue, that of data representation. Some of the approaches mentioned above operate on the data as measured (or perhaps pretreated in some way), such as (typically) morphological approaches or deformable models. Others are applied in a *feature vector* space. The division may be somewhat ambiguous — for example an MRF method may use both data and an ‘edge map’. It is possible for the same information to be stored as a *feature* or computed locally (e.g. local variance), in which case the information content is the same. In this way the segmentation algorithm and the representation may trade complexity. This noted, we may address them as separate, but related, concerns.

We believe that the best approach for the particular segmentation problems described here is to address it on two fronts. First, a multiscale analysis that will produce a data representation that embodies the robustness to changing orientations, and captures all of the relevant tissue information. This (redundant) space will become the so-called feature vectors that allow discrimination of texture (with a suitable metric). Secondly, an algorithm in the spirit of MRF approaches will optimize the segmentation of these vectors. Preliminary results are promising.

### 3 Data representations and types of textures

An image in its modern digital realization is an approximation of what we perceive as the original image. A widely accepted model for images postulates that an image is a function defined on a continuum, e.g.  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , has finite energy and bounded support (spatial extent). In the digitization process we create a finite array (e.g.  $512^2$  or  $512^3$ ) of numerical values to represent each image. Therefore, a digitized image is a function with a finite set of values defined on  $\mathbb{R}^2$  or  $\mathbb{R}^3$  with compact support, which is constant on a square/cube (voxel) with side length 1 centered at each point of the sampling lattice  $\mathbb{Z}^2$  or  $\mathbb{Z}^3$ . This model has been well-established in the literature, see.<sup>15</sup> Since we are interested only in 3D-images we henceforth fix the dimension of the domain of images and of the sampling lattice to three. In this sense images can be considered as functions in  $L^2(\mathbb{R}^3)$ . Our goal is to develop a theory for texture segmentation of 3D-images that is invariant under Euclidean rigid motions and is based on multiscale representations. We choose multiscale representation because textural properties manifest across scales.

Texture segmentation/identification algorithms may rely first on (often multi-scale) filtering techniques, which to every image  $f$  in  $L^2(\mathbb{R}^3)$  and each pixel  $z \in \mathbb{R}^3$  associate a ‘feature’ vector belonging to a fixed finite dimensional space  $V$ , endowed with a norm arising from an inner product  $\|\cdot\|$  on  $V$ . The vector space of all  $V$ -valued functions  $h$  defined on  $\mathbb{R}^3$  such that  $\|h(\cdot)\|$  is square integrable on  $\mathbb{R}^3$  is denoted by  $H$ . Thus,  $\|h\| = (\int_{\mathbb{R}^3} \|h(z)\|^2 dz)^{1/2}$ . To create features for each image  $f$  we use a linear transformation  $A$  associating to each  $f$  the  $V$ -valued function  $Af$  defined on  $\mathbb{R}^3$ . This linear transformation must be bounded. The values  $Af(z)$  are the *feature vectors* corresponding to  $f$ . Thus,  $V$  is also referred to as the *feature vector space* and  $A$  as the *feature map*. The boundedness hypothesis corresponds to a robustness property of the feature map  $A$ : small perturbations of pixel intensities for the image  $f$  generate small perturbations of the features  $Af$ .

We want to construct texture identification and associated segmentation algorithms which remain invariant under rigid motions of the image. Let  $GL(V)$  be the group of invertible linear maps of the finite-dimensional space  $V$  into itself. The group  $G$  of rigid motions  $\mathcal{R}$  of  $\mathbb{R}^3$  is generated by translations and rotations and acts naturally on images  $f$  by  $\mathcal{R}f(z) = f(Rz)$ . For translations we shall note  $T_a f(z) := f(z - a)$ , for  $a \in \mathbb{R}^3$ .

**DEFINITION 3.1.** Consider a bounded linear feature mapping  $A$  as above, generating for each image  $f$  and each voxel  $z$  in  $\mathbb{R}^3$  a feature vector  $Af(z)$  in the Euclidean space  $V$ . We shall say that  $A$  is a steerable feature mapping

if there is a mapping  $U$  from the group  $G$  of rigid motions into the general linear group  $GL(V)$  of  $V$  such that for each rigid motion  $\mathcal{R}$  of  $\mathbb{R}^3$ , the invertible transformation  $U(\mathcal{R})$  verifies

$$A[\mathcal{R}f](z) = U(\mathcal{R}) [Af(\mathcal{R}(z))] \quad z \in \mathbb{R}^3.$$

It can easily be shown that the mapping  $U$  must be a group homomorphism of the rigid motion group  $G$  into  $GL(V)$ , i.e.  $U(\mathcal{R}_1\mathcal{R}_2) = U(\mathcal{R}_1)U(\mathcal{R}_2)$  for all rigid motions  $\mathcal{R}_1$  and  $\mathcal{R}_2$ . However,  $U(G)$  is not abelian.

A very interesting class of steerable feature maps will arise from the constraint  $U(\mathcal{R}) = I_V$ , the identity on  $V$ , whenever  $\mathcal{R}$  is a translation. In this case the feature mapping  $A$  can be called *translation-invariant*. There are many examples of feature maps that are translation-invariant but most of them fail to be steerable: Suppose that the images  $f$  we consider have  $N^3$  values. If  $W := [-\frac{1}{2}, \frac{1}{2}]^3$  we define the feature map in the following way:

$$Af(z) := \int_{\mathbb{R}^3} f(z-s)\chi_W(s)ds, \quad z \in \mathbb{R}^3, \quad (1)$$

where  $\chi_W$  is the indicator (characteristic) function of  $W$ . Obviously

$$Af(z) = \langle f, T_z\chi_W \rangle = \int_{\mathbb{R}^3} \hat{f}(\xi) \overline{\widehat{\chi_W}(\xi)} e^{-2\pi i \xi z} d\xi.$$

This together with the fact that  $\hat{f}(\widehat{\chi_W}) \in L^1 \cap L^2$  show that  $A$  is a bounded operator. Obviously  $A$  is the feature map ‘intensity values’ and  $V$  is 1-dimensional.  $A$  is translation-invariant, but it is not steerable: Take the image  $\chi_W$  and rotate it by  $\pi/4$  keeping the  $z$ -axis fixed. Then, it is not hard to check that there is no non-zero number  $u_{\pi/4}$  satisfying

$$A[\mathcal{R}_{\pi/4}f](z) = u_{\pi/4} [Af(\mathcal{R}_{\pi/4}(z))] \quad \text{for all } z \in \mathbb{R}^3$$

as Definition 3.1 demands. The feature map defined in (1) can be defined instead of  $\chi_W$  for any square integrable, symmetric with respect to all three axes, window function, whose Fourier transform is bounded. In the light of the preceding discussion, an interesting mathematical problem is to *characterize the translation-invariant steerable feature maps*. The representation  $U$  defined according to Definition 3.1 is completely defined by its restriction to the Euclidean rotations centered at the origin. Specific examples of steerable multiscale families of wavelets have been given and used in image analysis. However, to the best of our knowledge no mathematical definition like 3.1 appears in the literature.

We are currently deriving a complete characterization of multiscale steerable feature maps which are also translation-invariant. Our formal steerability requirement is essential to construct computable discriminating local characteristics of textures, which will remain invariant when the texture image is arbitrarily rotated or translated in 3D. The focus of the present paper is tissue segmentation in medical images based on the use of a multi-scale translation-invariant steerable feature map, constructed by tensor products of Gaussian radial functions with spherical harmonics. This map is defined in Eq. (2) below.

A precursor of a steerable feature map results from a multiresolution analysis and is thoroughly studied in<sup>16</sup> and first proposed in.<sup>17,18</sup> We refer to this structure as First Generation Isotropic Multiresolution Analysis. Suppose that  $\phi$  is a refinable function such that  $\hat{\phi}$  is smooth and

- vanishes outside the ball centered at the origin with radius  $r < 1/2$
- $\hat{\phi}(\xi) = 1$  for all  $\xi$  in the ball  $B(0, \frac{1}{4})$

Define,  $\psi$  to be the inverse Fourier transform of  $\hat{\phi} - \hat{\phi}(2\cdot)$ .

The feature map  $\mathcal{A}$  we used was arrived at by means of a fast (Isotropic) 3D wavelet transform<sup>17</sup> at the every point of a regular  $N^3$ -lattice. This follows from a multiresolution analysis, thus the use of a refinable function  $\phi$

to generate this structure. Using the nomenclature introduced earlier in this section,  $A_0$  can be modeled by a continuous wavelet transform as follows:

$$\mathcal{A}f(z) = \left( \langle f, 8T_{4z}\phi(\cdot/4) \rangle, \langle f, 2^{3/2}T_{2z}\psi(\cdot/2) \rangle, \langle f, T_z\psi \rangle \right) \quad z \in \mathbb{R}^3.$$

Notice that  $V$  is three dimensional and that both  $\phi$  and  $\psi$  are radial functions. The first coordinate of  $\mathcal{A}f(z)$  corresponds to the LL-subband, while the second corresponds to the HL-subband. The high pass of  $f$  is not downsampled as indicated by the third coordinate of the feature vectors  $\mathcal{A}f(z)$ . The continuity of  $\mathcal{A}$  follows easily from the boundedness of  $\hat{\phi}$ . However, the feature map  $\mathcal{A}$  is not steerable: steerability fails because the condition of Definition 3.1 is not satisfied by all translations. On the other hand the same condition is valid for all rotations because  $\phi$  and  $\psi$  are radial.

The next step is to develop the steerable family we use for tissue segmentation in this paper. Take  $\sigma_j > 0$ ,  $j = 1, 2, \dots, K$  and the spherical harmonics  $Y_{m,n}$  with  $0 \leq m \leq M$  and  $n$  appropriately determined. The index  $m$  determines the degree of the spherical harmonic, with  $m = 0$  allowing only  $n = 0$  and  $Y_{0,0} = 1$ . For every  $m > 0$ , the set  $\{Y_{m,n} : 1 \leq n \leq d_m\}$ , where  $d_m$  is equal to 2 in the case of  $\mathbb{R}^2$  or  $2m + 1$  in the case of  $\mathbb{R}^3$ , is an orthonormal basis for the restrictions to the unit sphere  $\mathbb{S}^{d-1}$  ( $d = 2, 3$ ) of all homogeneous harmonic polynomials of degree  $m$  (see Chapter 4<sup>19</sup> or Chapter 9<sup>20</sup>).

Let  $G_{\sigma_j}$  be the Gaussian

$$G_{\sigma_j}(x) = e^{-\frac{\|x\|^2}{\sigma_j}} \quad x \in \mathbb{R}^3$$

and  $g_{\sigma_j}(x) = G_{\sigma_j}(x) - G_{\sigma_j}(4x)$ . Next, define

$$h_{j,m,n}(r, \omega) = g_{\sigma_j}(r)Y_{m,n}(\omega) \quad r > 0, \quad \omega \in \mathbb{S}^{d-1},$$

where  $d = 2, 3$ . The mapping  $A$  defined by

$$\mathcal{A}f(z) := \{ \langle f, T_z h_{j,m,n} \rangle : j = 1, 2, \dots, K, \quad 0 \leq m \leq M, \quad 1 \leq n \leq d_m \} \quad z \in \mathbb{R}^d \quad (2)$$

is the feature map we use to segment medical images in different tissue types in this paper. The feature space has dimension  $\sum_{m=0}^M K d_m$ . First, we must establish that  $A$  is bounded. Indeed,

$$\begin{aligned} \|\mathcal{A}f\|^2 &= \sum_{j,m,n} \int_{\mathbb{R}^3} |\langle T_{-z}f, h_{j,m,n} \rangle|^2 dz = \sum_{j,m,n} \int_{\mathbb{R}^3} \left| \int_{\mathbb{R}^3} \hat{f}(\xi) \overline{\widehat{h_{j,m,n}}(\xi)} e^{-2\pi i \xi z} d\xi \right|^2 dz \\ &= \sum_{j,m,n} \int_{\mathbb{R}^3} \left| \left[ \widehat{\widehat{f} \widehat{h_{j,m,n}}} \right]^\wedge(z) \right|^2 dz = \sum_{j,m,n} \|\widehat{f} \widehat{h_{j,m,n}}\|^2. \end{aligned}$$

Next we show that  $A$  is steerable and translation-invariant. While the latter property is obvious, the steerability with respect to rotations requires some discussion. First notice that rotations can be regarded as elements of the group  $SO(d)$ . Spherical harmonics transform to other spherical harmonics of the same degree under rotations. Indeed, if  $\mathcal{R}$  is a rotation, and  $p$  is a homogeneous harmonic polynomial of degree  $m$ , then  $\mathcal{R}p$  is again a homogeneous harmonic polynomial of the same degree. On the other hand, if  $f$  is a harmonic function, i.e. the Laplacian  $\Delta$  of  $f$  vanishes, then  $\mathcal{R}f$  is also harmonic (to establish this result calculate  $\nabla(\mathcal{R}f)(x)$ , then use the fact  $\Delta f(x) = \nabla f(x)(\nabla f(x))^T$ ).

Now it becomes apparent that if  $R$  is a rotation acting on  $\mathbb{R}^d$ , then  $Y_{m,n}(R \cdot)$  must be expressed in the terms of the orthonormal basis  $\{Y_{m,n} : 1 \leq n \leq d_m\}$ . What is challenging is to find the unique Fourier coefficients of this expression. Thanks to representation theory these coefficients can be explicitly calculated (see,<sup>20</sup> Section 9.4). We will not give any details of this calculation here. We simply conclude that if  $\mathcal{R}$  is a rotation, and  $m \geq 0$ , then there exists a unitary  $d_m \times d_m$  matrix, say  $\mathcal{M}(m, \mathcal{R})$  such that the Fourier coefficients of the expansion of  $\mathcal{R}(Y_{m,n})$  with respect to the orthonormal basis  $\{Y_{m,n} : 1 \leq n \leq d_m\}$  are the elements of the  $n$ -th column of the matrix  $\mathcal{M}(m, \mathcal{R})$ .

Let us see what happens to any one of the coordinates of the feature vector  $Af(z)$  if an arbitrary 2D- or 3D-image  $f$  is rotated by  $\mathcal{R}$  and  $A$  is the feature map defined in (2):

$$\langle \mathcal{R}f, T_z h_{j,m,n} \rangle = \langle f, \mathcal{R}^{-1} T_z \mathcal{R} \mathcal{R}^{-1}(h_{j,m,n}) \rangle = \langle f, T_{Rz} \mathcal{R}^{-1} h_{j,m,n} \rangle.$$

Now,

$$\mathcal{R}^{-1} h_{j,m,n}(r, \omega) = g_{\sigma_j}(r) Y_{m,n}(R\omega) = g_{\sigma_j}(r) \sum_{l=1}^{d_m} \mathcal{M}(m, \mathcal{R}^{-1})_{l,n} Y_{m,l}(\omega) = \sum_{l=1}^{d_m} \mathcal{M}(m, \mathcal{R}^{-1})_{l,n} h_{j,m,l}(r, \omega)$$

for all  $r > 0$  and  $\omega \in \mathbb{S}^{d-1}$ . Thus,

$$\langle \mathcal{R}f, T_z h_{j,m,n} \rangle = \sum_{l=1}^{d_m} \mathcal{M}(m, \mathcal{R}^{-1})_{l,n} \langle f, T_{Rz} h_{j,m,l} \rangle \quad \text{for all } j.$$

So, we conclude

$$A[\mathcal{R}f](z) = \mathcal{M}(m, \mathcal{R}^{-1}) Af(Rz) \quad \text{for all } z \in \mathbb{R}^d$$

which completes the proof of the steerability of this particular feature map.

We will close the present section with an remark: *What makes  $A$  steerable is its additional redundancy in translations. However, the capabilities of  $A$  in texture segmentation are better than those of  $\mathcal{A}$ , because locally the atoms of  $A$  (the tensor products of a radial Gaussian with spherical harmonics) can better match oscillating patterns in the angular variables, whereas  $\mathcal{A}$  simply ‘locally low-passes’ in the angular variables because its atoms are locally constant in the angular variables.*

## 4 Segmentations

Consider a steerable feature representation  $A$  as described in the previous paragraph. For each image  $f$  and pixel  $z \in \mathbb{R}^3$ , the local image features  $Af(z)$  is a vector in the euclidean space  $V$ .

**DEFINITION 4.1.** *Classically, unsupervised segmentation of an image into  $N$  regions  $\text{Reg}_1, \dots, \text{Reg}_N$  is a labeling  $\mathcal{L}$  of the image pixels with values  $1, \dots, N$  so that whenever  $\mathcal{L}(z) = j$ , the pixel  $z$  belongs to region  $\text{Reg}_j$ . Useful labellings (segmentations)  $\mathcal{L}$  must satisfy well known qualitative requirements, including an optimal compatibility with the feature mapping  $A$ : this essentially means that for neighboring pixels  $y$  and  $z$ , the feature vectors  $Af(z)$  and  $Af(y)$  must be very similar whenever  $y$  and  $z$  are in the same region, and must be very different whenever  $y$  and  $z$  are not in the same region.*

To formalize this requirement we select a ‘texture distance’  $D$  resulting from an inner product with which we endow the feature space  $V$ . Since we want texture segmentation/identification to remain robust to unknown rigid motions  $R$  of the image  $f$ , this naturally lead us to pursue the construction of *robust texture distances*  $D$  on the space  $V$  satisfying  $D(v, w) = D(v, U(R)w)$  for all rigid motions  $R$  in the group  $G$ .

Formally we will define a *texture type* to be any compact subset  $T$  of the feature space  $V$ . In ‘supervised segmentation’ tasks, concrete texture types of interest are most often defined by medical experts who use visual or histological identification of tissue types on specific image patches. This enables the generation of disjoint compact subsets  $T_1, \dots, T_p$  of  $V$  representing the distinct ‘texture types’ identified.

The goal of a *supervised segmentation* algorithm is to generate a segmentation  $\mathcal{L}$ , which assigns to each pixel  $z$  a *label*  $\mathcal{L}(z) \in 1, \dots, p$  defining the texture type of pixel  $z$ . The mapping  $\mathcal{L}(z) = j$  identifies the point  $z$  with ‘texture type’  $T_j$ . Supervision may be provided in picking the number of labels,  $p$ , in defining characteristics of ‘texture types’, or both.

Moreover we want the texture type identification  $\mathcal{L}(z) = j$  at pixel  $z$  to remain the same if we first rotate the underlying image  $f$  and then segment it. This is a strong requirement on acceptable segmentation algorithms, and will be formalized elsewhere. But this imposes the following requirement:

*Rigid motion invariance of texture types:* For each  $j = 1, \dots, p$  the subset  $T_j$  of  $V$  must be invariant by every  $U(R)$  in  $GL(V)$ , where  $R$  is an arbitrary rigid motion of  $\mathbb{R}^3$  and  $U$  is the homomorphism associated to the steerable feature map  $A$ .

#### 4.1 Segmentation Algorithm

As previously noted, the algorithmic approach can be considered in two parts — the underlying multiscale representation, and the segmentation algorithm that operates in this space. The previous section describes our ongoing approach to choosing the appropriate feature vector space and distance. In moving beyond our First Generation IMRA based segmentation techniques, we are developing a novel segmentation algorithm in the spirit of Markov Random Field approaches utilizing the steerable translation-invariant feature map  $A$  defined by Eq. (2).

We are working in three dimensions; there are significant practical constraints on the amount of computation that will remain tractable for the problem domain. For this reason, some of the design constraints are as much computational as mathematical. The quality of any potential segmentation  $\mathcal{L}$  is quantified by a *cost function* usually referred to in the MRF context as an *energy function*,  $\mathcal{E}(\mathcal{L})$ . Assigning a positive energy function to the candidate labeling function  $\mathcal{L}(z)$  for  $z \in \mathbb{R}^3$ . Finding a good segmentation  $\mathcal{L}$  becomes the problem of minimizing the energy  $\mathcal{E}(\mathcal{L})$ .

To remain computationally tractable, optimization was done by multi-scale non-stochastic gradient descent with initial texture types determined by a *k-means* on the feature space  $V$ . One cannot expect global optimality, but good results are possible while achieving computational efficiency. Our ongoing work is exploring the capabilities of this class of algorithms, and of the underlying representations. The preliminary results reported in §5 are promising.

As previously noted, the energy of a segmentation  $\mathcal{L}$  requires the definition of a ‘texture distance’  $D$  placed on the feature vector space  $V$ . We are currently comparing various appropriate metrics; the experiments reported in this paper simply use a Euclidean distance on the feature vectors, but our final goal is to implement distances  $D$  between textures incorporating the ideas of robustness to rotations.

Each voxel  $z \in \mathbb{R}^3$  has a label  $\mathcal{L}(z)$  in  $1, \dots, N$ . Each voxel also has a localized energy  $\mathcal{E}_z$ , which is the sum of four components:  $\mathcal{E}_z = \sum_{k=1}^4 V_k(z)$ . Each component will typically take contributions from neighboring pixels (i.e. the elements of a neighborhood  $\mathcal{N}(z)$ ), so

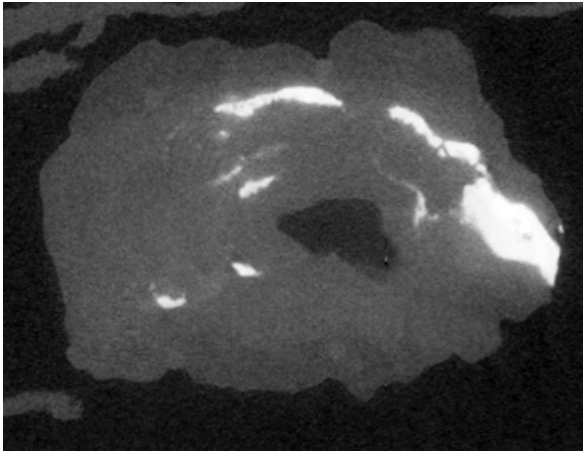
$$V_k(z) = \sum_{z' \in \mathcal{N}(z)} \alpha_k f_k(d(z, z'), \mathcal{L}(z), \mathcal{L}(z')) .$$

Here  $\alpha_k$  is a parameter and  $f_k$  a precise function of the distance between feature vectors at sites  $z$  and  $z'$ , and the labels of both pixels.

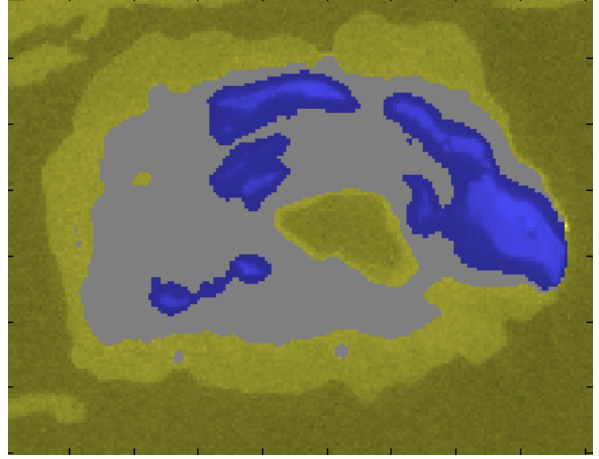
The four components of the energy function tend to enforce separate crucial qualitative aspects of the segmentation  $L$ :

1. *Texture discontinuities must lie on region boundaries*: local energy  $V_1$  will force region boundaries to coincide with strong discontinuities in the feature vector space.
2. *Regional homogeneity of textures*: local energy  $V_2$  ensures that when voxels  $z$  and  $z'$  lie within the same region, and are close to each other, then the distance between their feature vectors  $Af(z)$  and  $Af(z')$  must be small.
3. *Label consistency*: local energy  $V_3$  penalizes drift across regions, i.e. voxels in the same region should have statistically similar feature vectors even when they are spatially far apart.
4. *Smoothness of region boundaries*: enforced by local energy term  $V_4$  (smoothness constraints).

The energy of the entire segmentation is the sum of the local energies  $\mathcal{E} = \sum_{z \in \mathbb{R}^3} \mathcal{E}_z$ . Optimizing this function results in a segmentation.



(a) input  $\mu$ CT data



(b) First Generation IMRA-based segmentation

Figure 2: Segmentation performed on  $\mu$ CT volume of arterial segments. (a) shows a single slice through the data volume (b) shows a colored representation of the segmentation approach based on the feature map  $\mathcal{A}$

## 5 Results and Commentary

The largest part of our empirical work to date has involved micro x-ray CT ( $\mu$ CT) volumes of arterial specimens. This data set and the results of our First Generation IMRA approach are discussed in detail in.<sup>21</sup>

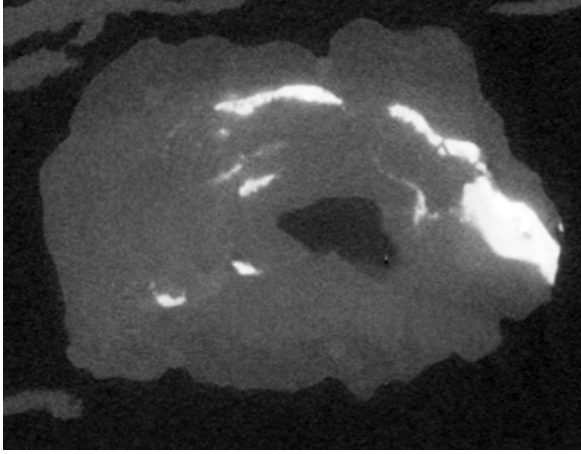
For our purposes here it suffices to show a characteristic result; Figure 2 (a) shows a slice of one calcified segment, while (b) gives a color overlay representation of the segmentation achieved with the feature map  $\mathcal{A}$  resulting from the First Generation IMRA. There are three classes — yellow, blue, and neutral gray. The yellow and blue are overlaid on actual data, and show regions of ‘abnormal’ tissue, as compared to the grayed out tissue. It should be noted that the overall purpose of that study was *not* to segment the images, and particularly the colored image shown here was not the output used. It does, however, show the characteristics of the segmentation part of our reconstruction algorithm. These results help to inform the new segmentation approaches we are constructing. Particular features to note are the lack of clear boundary on the arterial wall and lumen, and sometimes overly generous boundaries surrounding calcifications.

By comparison, Figure 3 again shows the original data, along with a segmentation result from the new algorithm. Again the color map is overlaid on the real data; there are only five labels (colors) and shading represents features of the underlying data. This result has greatly improved the boundary definition of arterial wall, lumen, and calcifications. It also finds two tissue types (red and light blue). There are, however, issues to be resolved in the consistency of these two tissue types.

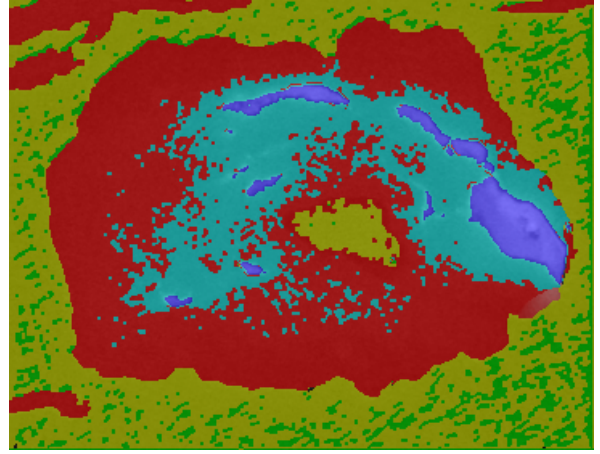
Figure 4 shows the results of decomposing the input data under our first generation IMRA scheme. This results in a *feature vector space* of dimension 3 (or 4 if original data is included). We can decompose further, but rapidly lose resolution, so for this application it is not helpful to do so (since relevant information about the textures is localized, and of relatively high frequency).

By comparison, the tensor product family described in Section 3 allows us much more freedom to design a feature vector space tuned to a particular applications needs. Figure 5 shows several results from applying various members of the tensor product family described previously; Gaussian and spherical harmonics in 2D. When,  $d = 2$  the orthonormal basis of the space of spherical harmonics of degree  $m$  consists of two functions:  $\cos(m \cdot)$  and  $\sin(m \cdot)$ . In Figure 5 we see outputs for two choices of the parameter  $\sigma_j$  (controlling the width of the Gaussian ‘window’) and for each of these, several choices of  $m$ , (controlling the ‘rate of local angular oscillation’). These are the parameters that were used to construct the feature vector space used in the segmentation result



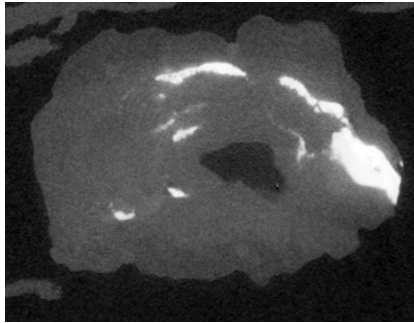


(a) input  $\mu$ CT data

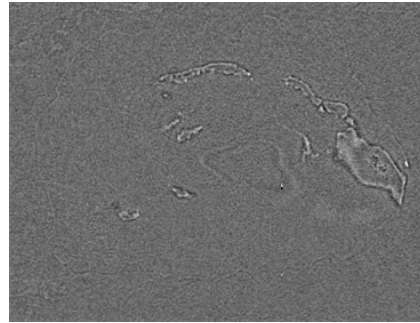


(b) new segmentation algorithm

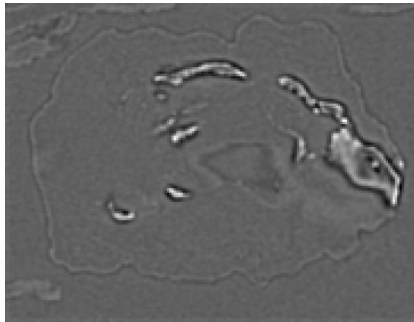
Figure 3: Segmentation performed on  $\mu$ CT volume of arterial segments. (a) shows a single slice through the data volume (b) shows a colored representation segmentation approach utilizing the steerable feature map defined by Eq. (2)



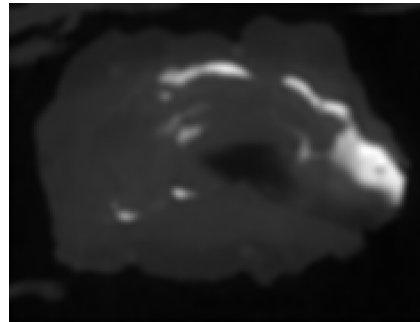
(a) input



(b) H

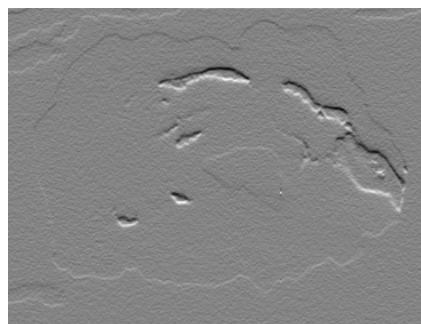


(c) LH

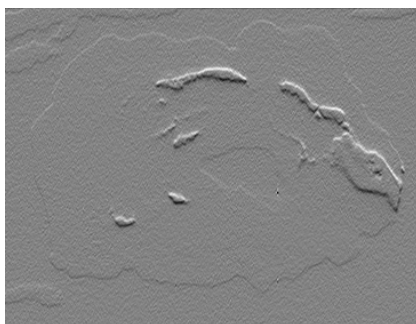


(d) LL

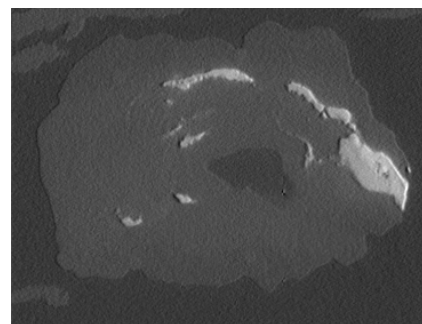
Figure 4: IMRA decomposition. Input data (a) is decomposed into (b) H and The L subband (not part of the representation) is further decomposed into (c) LH and (d) LL subbands.



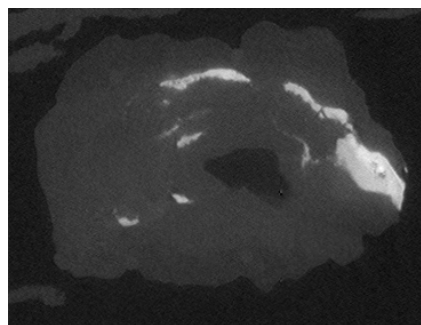
(a)  $\sigma = 0.4, n = 1, \sin$



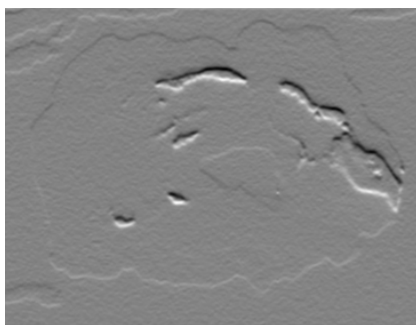
(b)  $\sigma = 0.4, n = 3, \sin$



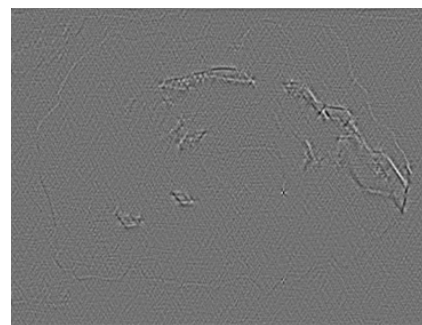
(c)  $\sigma = 0.4, n = 1, \cos$



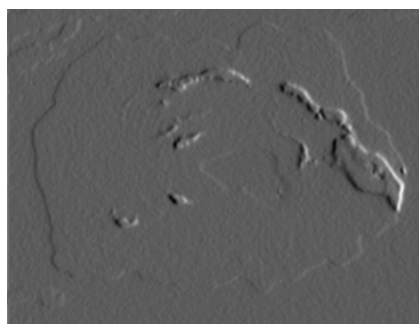
(d)  $\sigma = 0.4, n = 3, \cos$



(e)  $\sigma = 1.2, n = 1, \sin$



(f)  $\sigma = 1.2, n = 3, \sin$



(g)  $\sigma = 1.2, n = 1, \cos$



(h)  $\sigma = 1.2, n = 3, \cos$

Figure 5: Gaussian ‘windows’ tensor product with spherical harmonics. Here we see the outputs of different choices of  $\sigma_j$  and  $m$  for the 2D-spherical harmonics.

showed in Figure 3. The parameters are in no sense optimal. Inspection here shows that there is apparently quite a bit of redundant information in some of the bands, so this can be improved upon.

Finally, Figure 6 shows a comparison between original data, histological ‘ground truth’, and the result of the new segmentation algorithm. This figure demonstrates some difficulties with the modality; the strong presence

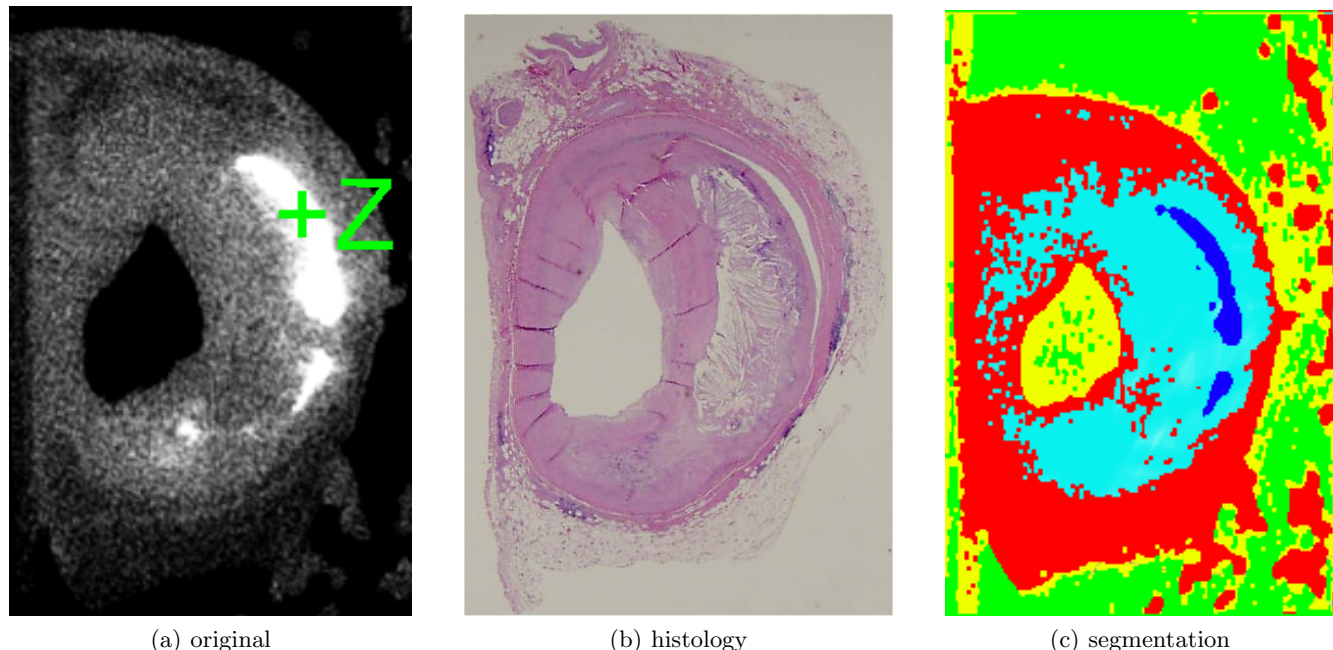


Figure 6: For this data set we have histological truth. Here we see the original data (a) histology slide (b), and a segmentation of comparable slice of the  $\mu$ CT data (registered by hand).

of calcium in the data (a) overshadows texture detail. In the histological slide (b) a lipid pool bordering the location of calcium (deep blue in (c)) is very apparent. The segmentation in (b) does not circumscribe this lipid pool tightly, which it is not visually apparent in the data either. The segmentation does give good boundaries on other features, and recognizes the density differences in the arterial wall, but it is not really pinning down the lipid. This is a very difficult section to segment well due to the overwhelming scattering effect of the calcium in the original data. The result shown is quite good, but we are convinced it can be improved with further study.

In conclusion, We have demonstrated our approach to medical image segmentation, with reference to both the first generation IMRA based approach and our ongoing work on multiscale representations and feature vector spaces, ‘texture distances’ and segmentation algorithms. Preliminary results given here demonstrate good performance for certain tasks, and our ongoing work will generalize these results.

## REFERENCES

1. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
2. L. D. Cohen and I. Cohen, "Finite element methods for active contour models and balloons for 2d and 3d images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 11, pp. 1131–1147, November 1993.
3. V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *ICCV*, 1995, pp. 694–699. [Online]. Available: [citeseer.ist.psu.edu/caselles95geodesic.html](http://citeseer.ist.psu.edu/caselles95geodesic.html)
4. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, 1984.
5. J. Besag, P. Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems (with discussion)," *Statistical Science*, vol. 10, pp. 3–66, 1995. [Online]. Available: [citeseer.ist.psu.edu/besag95bayesian.html](http://citeseer.ist.psu.edu/besag95bayesian.html)
6. J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statistical Society B*, vol. 36, no. 192–225, 1974.
7. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988. [Online]. Available: [citeseer.ist.psu.edu/osher88fronts.html](http://citeseer.ist.psu.edu/osher88fronts.html)
8. R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 2, pp. 158–175, 1995. [Online]. Available: [citeseer.ist.psu.edu/article/malladi95shape.html](http://citeseer.ist.psu.edu/article/malladi95shape.html)
9. J. A. Sethian, *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.
10. A. Mitiche and S. Hicham, "Optical flow 3d segmentation and interpretation: A variational method with active curve evolution and level sets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 28, no. 11, pp. 1818 – 1829, 2006.
11. B. Parker and F. Dagan, "Graph-based mumford-shah segmentation of dynamic pet with application to input function estimation," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 1, pp. 79–89, 2005.
12. J. Smaarabandu, R. Acharya, and P. Cheng, *Analysis and Presentation of Three Dimensional Data Sets*. Berlin: Springer Verlag, 1993.
13. J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, 1982.
14. S. Beucher and F. Meyer, "The morphological approach of segmentation: the watershed transformation," in *Mathematical Morphology in Image Processing*, E. Dougherty, Ed. CRC Press, 1993, ch. 12, pp. 433–481.
15. D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure and Appl. Math.*, vol. 42, pp. 577–685, 1989.
16. S. Alexander, S. Baid, S. Jain, M. Papadakis, and J. Romero, "Isotropic multiresolution analysis: geometry and analytic properties," in progress.
17. B. Bodmann and M. P. et. al., "Frame isotropic multiresolution analysis for micro CT scans of coronary arteries," in *Wavelets XI*, M. Papadakis, A. Laine, and M. Unser, Eds., vol. 5914. SPIE, 2005, pp. 59141O/1–12.
18. M. Papadakis, G. Gogoshin, I. Kakadiaris, D. Kouri, and D. Hoffman, "Non-separable radial frame multiresolution analysis in multidimensions," *Numer. Function. Anal. Optimization*, vol. 24, pp. 907–928, 2003.
19. E. Stein and G. Weiss, *Introduction to Fourier Analysis in Euclidean spaces*. Princeton, NJ: Princeton University Press, 1971.
20. G. Chirikjian and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotation and Motion Groups*. CRC Press, 2000, electronic resource.
21. S. Gertz, B. Bodmann, D. Vela, M. Papadakis, and et.al., "Three-dimensional isotropic wavelets for post-acquisitional extraction of latent images of atherosclerotic plaque components from micro-computed tomography of human coronary arteries," 2007, in print.