

MIT Open Access Articles

Hierarchical 3D diffusion wavelet shape priors

The MIT Faculty has made this article openly available. *Please share* how this access benefits you. Your story matters.

Citation: Essafi, S., G. Langs, and N. Paragios. "Hierarchical 3D diffusion wavelet shape priors." Computer Vision, 2009 IEEE 12th International Conference on. 2009. 1717-1724. © 2010 Institute of Electrical and Electronics Engineers.

As Published: http://dx.doi.org/10.1109/ICCV.2009.5459385

Publisher: Institute of Electrical and Electronics Engineers

Persistent URL: http://hdl.handle.net/1721.1/59305

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Hierarchical 3D Diffusion Wavelet Shape Priors

Salma Essafi * Laboratoire MAS, Ecole Centrale Paris, France Galen Group, INRIA Saclay, France Georg Langs CIR Lab, Medical University of Vienna, Austria CSAIL, MIT, USA

salma.essafi@ecp.fr

langs@csail.mit.edu

Nikos Paragios Laboratoire MAS, Ecole Centrale Paris, France Galen Group, INRIA Saclay, France

nikos.paragios@ecp.fr

Abstract

In this paper, we propose a novel representation of prior knowledge for image segmentation, using diffusion wavelets that can reflect arbitrary continuous interdependencies in shape data. The application of diffusion wavelets has, so far, largely been confined to signal processing. In our approach, and in contrast to state-of-the-art methods, we optimize the coefficients, the number and the position of landmarks, and the object topology - the domain on which the wavelets are defined - during the model learning phase, in a coarse-to-fine manner. The resulting paradigm supports hierarchies both in the model and the search space, can encode complex geometric and photometric dependencies of the structure of interest, and can deal with arbitrary topologies. We report results on two challenging medical data sets, that illustrate the impact of the soft parameterization and the potential of the diffusion operator.

1. Introduction

Segmentation is a fundamental problem in computer vision and medical imaging. It consists of automatically partitioning data into a number of disjoint regions according to their appearance properties. State of the art methods involve model-free and model-based ones. Model-free methods make no assumption on the geometric properties of the region of interest. Model-based methods introduce certain assumptions on the space of allowable solutions - *priors*. Approaches relying on a prior are useful in the context of medical image analysis where variations of anatomical structures are constrained by the anatomy, while at the same time pose and view-point variation can be dealt with. Modeling shape variation is a well studied problem, with two critical components, the choice of shape representation and the construction of the prior manifold. Point distribution or landmark-based models [8], implicit representations [22], and spherical wavelet representations [21] are examples of such shape and surface representations.

Given the shape representation, the prior manifold can either be a subspace or a probability density function. In the first case, the space of solutions is often represented using a linear combination of a set of basis functions modeling the variations of the training examples. Linear sub-spaces, determined either through Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), or Non Negative Matrix Factorization (NNMF), are methods being used to determine these subspaces. Toward dealing with high amounts of training data, numerous provisions were considered such as Kernel-PCA prior [9]. In the second case, non parametric densities [25] as well as manifold learning and embedding [3] were considered.

Once the manifold has been determined, one can either utilize knowledge-based methods where the solution should live on the manifold, like active shape and appearance models [8, 7]. The main strength of these methods is robustness, while their main limitation is that they can only account for variations being observed in the training set. Alternatively one can use more flexible methods combining prior with data terms aiming to maximize the support from the image, while minimizing the distance from the learned manifold like for example Mumford-Shah functional shape priors [9]. These methods can deal reasonably well with variations not being observed in the training set, if image support is adequate, but first require a significant number of training samples toward learning the density. They are also quite sensitive to the initial conditions and are computationally expensive.

^{*}This work was supported by Association Française contre les Myopathies (AFM: http://www.afm-france.org) under the DTI-MUSCLE project.

Ideally one would like to optimize both the shape representation and the compactness of the manifold. In this paper we propose a novel hierarchical surface representation that can encode shape variation of structures with arbitrary topology, in a compact statistical manifold. The shape representation is based on a finite set of landmarks that exhibit significant differentiation to the background on different examples of the anatomical structure. In contrast to conventional shape modeling approaches that rely on a pre-defined topology, and parameterize the surface of an object with regard to a manifold like a sphere, the proposed method learns the appropriate topology from the training data, and uses a corresponding shape representation based on diffusion wavelets to model its variation. The approach encodes hierarchies, and can deal with complex and soft connectivity properties of objects by encoding their interdependencies with a diffusion kernel [5] that defines the domain of the wavelets.

The topology, determining the wavelet domain, is learned from the training data instead of using a priori choices and the shape variation is represented by means of the corresponding diffusion wavelets [6]. To build a model, statistical learning of the variation of the wavelet coefficients at different levels of the hierarchy is performed. Due to the power of the basis function representation, dimensionality reduction techniques using the orthomax criterion [15] lead to a very compact representation of the manifold.

During model search these constraints are used to infer landmark positions from volume data guided by a data term, and the diffusion wavelet shape constraint. The optimal solution given the constraint is determined by a simple projection to the linear manifold. Promising results both in terms of reconstruction error and segmentation demonstrate the extreme potentials of our method using the challenging case of T1 calf images.

There are several works related to this approach. In [10] the idea to build a hierarchical active shape models of 2-D anatomical objects using 1-D wavelets, which are then used for shape based image segmentation was explored. An exciting extension was proposed in [21], where spherical wavelets characterize shape variation in a local fashion in both space and frequency on a spherical domain. More recently in [26] the idea of using diffusion wavelets as a shape descriptor for the matching of 3D shapes was considered. The work proposed in this paper differs both in terms of method and scope: in contrary to [26], we are building a generative model of shape variation, which adapts to the topology of a set training examples, and use it for the segmentation and reconstruction in new volume data. An important advantage of the proposed method is the use of orthomax to obtain a sparse representation of the shape variation in the coefficient space. Hierarchical shape models based on diffusion wavelets that adapt to arbitrary topolo-

1718

gies have not yet been published, and their use in computer vision and medical image analysis has not been described. With respect to segmentation, spherical wavelets are a subcase of diffusion wavelets confined to a spherical manifold. With respect to registration approaches, the prime difference is that we are building a generative model of shape variation, and use it for segmentation while at the same time using the orthomax criterion to obtain an optimal subdivision of the shape parameterization.

The remainder of the paper is organized as follows: in Section. 2, the shape representation using diffusion wavelets is presented, and in Section. 3 we focus on the manifold construction and the inference in new data. In Section. 4 we present the context of the work and report experimental results and a quantitative validation. Finally Section. 5 concludes with a discussion and future directions.

2. Shape Representation

We parameterize the shape variation observed in the training data by means of diffusion wavelets. The topology - the domain on which we define the wavelet representation - of the structure that is modeled is defined by a diffusion kernel. It can be viewed as a generalization of standard parameterizations, e.g.: the diffusion kernel for a triangulated spherical surface would be the adjacency matrix weighted by the mutual distances. Defining the topology by a diffusion kernel instead of a fixed genus-0 manifold allows us to incorporate, and even to learn, complex interaction patterns observed in the training data, and use them to build an efficient shape variation prior. In the following we will first outline the basics of diffusion wavelets [6], then explain the associated shape variation representation, and finally detail how the orthomax principle [15] can be used to separate coherent sub-regions of the shape.

2.1. Diffusion Wavelets

Wavelets are a robust mathematical tool for the hierarchical decomposition of functions. The theory is described extensively in e.g. [19]. The decomposition allows for a representation in terms of a coarse overall shape, that is enriched by details in a coarse to fine hierarchy. They provide an elegant technique for representing detail levels regardless of the interest function type (e.g. images, curves, surfaces). Their major strengths are the compact support of basis functions as well ad the inherently hierarchical representation. The domain upon which the wavelet hierarchy is defined is of prime importance for their representative power.

In this paper we will use wavelets to represent the variation of shapes. However, instead of relying on a pre-defined manifold (e.g. a sphere) we will learn the topology of the wavelet domain from the training data, and will encode it in a diffusion kernel. The kernel allows us to learn and define arbitrary wavelet hierarchies, and thus to make optimal use of the training data. The wavelet representation used in our paper is based on *diffusion wavelets* proposed in [6].

2.2. A Diffusion Operator Reflecting The Topology

We represent the shapes by a finite set of landmarks. For m landmark positions, $\mathbf{V}_i = \langle \mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_m^i, \rangle$, are known in n training images $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N$. That is, our shape knowledge comprises $\mathcal{V} = \{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N\}$, where $\mathbf{x}_j^i \in \mathbb{R}^d$, and we call $\mathbf{V}_i \in \mathbb{R}^{dm}$ a shape.

Since we are only interested in the non-rigid deformation, all anatomical shapes are aligned by Procrustes analysis [18], which produces the series of examples V_i^p , from which we compute the mean shape \bar{V}^p . After the registration, we can represent the shapes by their deviation S_i from the mean shape, $\bar{V}^p = \frac{1}{N} \sum_{i=1}^{N} V_i^p$, where $S_i = V_i^p - \bar{V}^p$.

Now we define a topology on the set of landmarks. The representation is based on a framework for multi-scale geometric graph analysis proposed in [5]. It applies the concept of diffusion to capture mutual relations between nodes in a Markov chain, and derive the global structure of the shape. Indeed diffusion maps provide a canonical representation of high-dimensional data. They allow us to encode spatial relations, or the behavior [17] of the shape training population. The structure is encoded in a diffusion operator $T \in \mathbb{R}^{m \times m}$. The association with diffusion wavelets enables us to directly use both the global and local properties of the data encoded in T.

Within our approach, we build the diffusion operator T on the set of points embedded in a metric space in two different manners, either using: [(a)] Their mutual distance in the mean shape, or [(b)] their joint modeling behavior. In the first case, and in order to build a matrix of graph weights for the points, we construct a local Gaussian kernel function centered at each point and then normalize the weight matrix through the symmetric Laplace-Beltrami to form the diffusion operator T. In the second case, when seeking modeling their joint behavior, we derive the diffusion operator by probing the behavior of small subsets of the landmark set, according to the method described in [17]. The resulting operator T reflects all pairwise relations or *neighborhoods* between individual points in the shape set.

2.3. Shape Variation Modeling with Diffusion Wavelets

Given this diffusion operator T defining the manifold, we use the corresponding hierarchical diffusion wavelets, to represent the shape variation. First we build a hierarchical wavelet structure, the *diffusion wavelet tree*: We call upon a general multi resolution construction for efficiently computing, representing and compressing T^{2j} , for j > 0. The latter are dyadic powers of T, and we use them as dilation operators to move from one level to the next. We can expect it to be easier to compress high orders of the diffusion operator as they are supposed to be low ranked. During the down-sampling process, and throughout a recursive sparse QR decomposition we obtain the orthonormal bases of scaling functions, $\Phi = {\phi_j}$, the wavelets W_j , and compressed representation of T^{2j} on ϕ_j , for j in the requested range. Giving K as maximum number of levels to compute, we obtain a representation of T^{2j} onto a basis ϕ_j , with $1 \le j \le K$ after K steps. For a detailed description of this construction we refer the reader to [6].

After building the diffusion wavelet tree Φ , we use it to represent the individual training shapes. We calculate the diffusion wavelet coefficient Γ on the deviation S_i from the mean of the aligned shapes, and obtain the following diffusion wavelet coefficients for an example S_i , $\Gamma_{S_i} = \Phi^{-1}S_i$. Thus, the shape can be reconstructed by:

$$V_i^p = \bar{V}^p + \Phi \Gamma_{S_i} \tag{1}$$

Once we have generated the diffusion wavelet coefficients for all training examples, we build a model of the variation by means of the orthomax criterion - which will be described in the upcoming section - at each level. In the lowest level the coefficients provide information for a coarse approximation, whereas localized variations are captured by the higher-level coefficients in the hierarchy. For each level j, with $(1 \le j \le K)$, we consider Γ_{levelj} (Eq. (2)) and perform principle component analysis to reduce the dimension of the coefficient representation for all coefficients scales.

$$\Gamma_{levelj} = \left\{ \Gamma_{S_i/level=j} \right\}_{i=1\dots N},\tag{2}$$

This results in the eigenvectors $\Sigma = {\sigma_j}_{j=1...K}$, the corresponding eigenvalues $\Lambda = {\lambda_j}_{j=1...K}$ of the covariance matrix of the diffusion wavelets coefficients at each level j, and the according coefficients $\Gamma_{level}_j^*$ that represent each training shape in this coordinate system.

Consequently in each level the coefficients would be expressed such as:

$$\Gamma_{levelj} = \overline{\Gamma}_{level_j} + \sigma_j \left(\sigma'_j \cdot \Gamma_{level_j^*} \right)$$
(3)

Based on the model parameters $\{\Lambda, \Sigma\}$ we can reconstruct a shape by first obtaining the diffusion wavelet coefficients $\Gamma_{S_{i}Rec}$ in each level, and then reconstructing the shape based on the diffusion wavelet tree:

$$V_i^p = \bar{V}^p + \Phi \Gamma_{S_i Rec} \tag{4}$$

This shape representation can now be considered for manifold learning. The idea will be to determine a manifold on the diffusion wavelet space using the distribution of coefficients of the training data.



Figure 1. Data reconstruction through (a) global PCA and (b) localized orthomax rotated modes of the DW coefficients models. The surfaces representing the $\pm 3sqrt(\lambda_i)$ are respectively colored in cyan, red and blue. Note the local deformation captured by the orthomax mode.

3. Prior Manifold Construction & Imagebased Inference

Conventional dimensionality reduction techniques like PCA, LDA, NNMF, statistical approximation methods like mixture models, EM, or recent spectral kernel methods like Locally Linear Embedding [23] or Laplacian Eigenmaps [1] can be considered. Given the ability of the diffusion model to capture relevant non-linear variations, we chose a simple dimensionality reduction technique, and a linear sub-space representation for our experiments. For further condensing the model while preserving its ability to capture the variations, we adopt the orthomax criterion [15].

3.1. Modeling Using the Orthomax Criterion

The orthomax criterion [24] allows to obtain a simple and compact hierarchical representation through a rotation of the model parameter system. We explore the varimax version [15] for optimizing sparsity corresponding to new variables being associated to localized variation modes. Orthomax rotations reflect a re-parameterization of the PCA space resulting in a simple basis.

Let \mathbf{R} be an orthonormal rotation matrix in \mathbb{R}^k where $\mathbf{R}_{i,j}$ represents \mathbf{R} elements, and where k implies the number of eigenvectors with the largest eigenvalues λ_i . Besides Σ denotes as previously in the paper the $p \times N$ eigenvectors matrix. The orthogonal orthomax rotation matrix \mathbf{R} is calculated as follow:

$$\boldsymbol{R} = \arg \max_{R} \left(\sum_{j=1}^{k} \sum_{i=1}^{p} (\Sigma \boldsymbol{R})_{ij}^{4} - \frac{1}{p} \sum_{j=1}^{k} \left(\sum_{i=1}^{p} (\Sigma \boldsymbol{R})_{ij}^{2} \right)^{2} \right),$$
(5)

One can easily notice in Fig.(1) that while the PCA modes demonstrate several spatially distributed effects within each mode, the varimax modes in the other hand show nicely isolated effects. Moreover in Fig.(2), we show the 'flattening' of the eigenvalue spectrum carried out

by the varimax rotation where the respective modes as well as variances are plotted. This simple, yet powerful modification of PCA enables us to optimize sparsity leading to localized modes of variation, which is more suitable for applications with sparse parameterizations like the often local pathological variations we are focusing on.

Once we succeeded to separate the data variation through the wavelet level space, we can then get the shape prior by projecting back the selected orthomax eigenvectors of the diffusion coefficients level into the right dimensions. Let γ_j denote the orthomax eigenvectors such as $\gamma_j = \mathbf{R}^{-1}\sigma_j$. Then Eq. (3) can be expressed with the orthomax components.

$$\Psi_{level}^{i*} = \begin{bmatrix} \Psi_{level_1}^{i*} \Psi_{level_2}^{i*} \dots \Psi_{level_K}^{i*} \end{bmatrix}$$
(6)

$$V_i^p = \bar{V}^p + \Phi. \left(\bar{\Psi}_{level_j} + \gamma_j \left(\gamma'_j. \Psi_{level}^{i*} \right) \right), \quad (7)$$

An overview of the model building process (including the representation component) is given in (Alg.1). The resulting model holds information about the diffusion wavelet tree, the orthomax components, and coefficient variation constraints $\langle \Phi, \mathbf{R}, \gamma \rangle$.

Algorithm 1 Multi-	Scale Represen	tation: Off line	Training
Input:			

T: Diffusion Operator,

 $(\mathbf{S}_i)_{i \in \text{Training Shapes}}$: Training landmark deviation from the mean shape

- 1. Construct the wavelet diffusion tree Φ [6].
- 2. Calculate Diffusion Wavelet Coefficient for the training shapes S_i .
- 3. For each level.
 - (a) Compute covariance matrix of every diffusion wavelet coefficient level.
 - (b) Calculate a model parameterization (basis, coefficients) based on the orthomax criterion.

Output: Eigenspace (modes, eigenvalues) and coefficients of the training examples.

The outcome of this process is an efficient shape representation as well as a compact manifold construction with respect to the allowable variations of the this representation. We first obtain a topology from the training data and encode it in a diffusion kernel, that defines a diffusion process [4] across the set of landmarks. It can either be based on their distance, or on their mutual dependencies observed in the training data. Given this kernel, we build a hierarchical wavelet representation of the shape variation. Finally we build a sparse model of the individual levels of this representation with help of the *orthomax* criterion. Keep in mind



Figure 2. Comparaison between PCA and Orthomax DW eigenvalues

that the main goal is to capture meaningful structures based on their behavior in the potentially very small training data set. This representation, along with the manifold can now be used for image based inference in a new example.

3.2. Image-based Inference

The search with the diffusion model representation, and appearance patch models $(P_i)_{i=1,...,N}$ for each landmark is performed in an iterative manner, starting from a coarse initialization obtained by atlas registration. The appearance model is based on a local texture patch model at the landmark positions. The model is built analogously to [12]. Similar to a standard shape model inference scheme, the landmarks positions in new data are estimated by an energy minimization involving both shape prior and appearance costs:

1. The landmark positions are updated according to a local appearance model. For each landmark the position with highest probability of the corresponding local texture patch being consistent with regard to the learnt texture model P_i is chosen within a neighborhood to of the current position estimate. We consider $P(\mathbf{x}_i^j)$ as the learned texture patch for the correct landmark position \mathbf{x}_i in the initial training volume. As for landmark positions in a local neighborhood \mathcal{N} , we have $C_i^j(\mathbf{x})$ as the correlation between the patch $p(\mathbf{x})$ and $P(\mathbf{x}_i^j)$ normalized within the neighborhood, i.e. $\int_{\mathbf{x} \in \mathcal{N}} C_i^j(\mathbf{x}) = 1$, then the image support is

$$\xi_i = mean_{j=1,\dots,n} \left(\frac{C_i^j(\mathbf{x})}{\int_{\mathbf{x} \in \mathcal{N} \setminus \mathbf{x}_i^j} C_i^j(\mathbf{x})} \right).$$
(8)

The image support is thus computed for every landmark in \mathcal{V} from the local appearance behavior at the corresponding positions in the training shape.

2. After fitting the shape to the image data, its variation is constrained by the diffusion wavelet shape variation model. The landmarks are projected into the orthomax coefficient space as described in Sec 3.1. The constraints learned during training are applied, and based on the resulting parameter values the shape is reconstructed.

This procedure is iterated while during each iteration, the corresponding \mathbf{V}' is reconstructed to re-estimate the shape. After convergence the final reconstruction \mathbf{V}' is an estimate of the true shape inferred from the data, and the prior model.

4. Experimental Results

We evaluate the algorithm on two challenging medical imaging applications to assess the performance of the method in terms of representation, manifold construction and knowledge-based segmentation. The first example is the segmentation of the left ventricle (LV) of the heart using computed tomography images, and the second the segmentation of calf muscles from T1 Magnetic Resonance Images. While in the first case, the performance of the extraction of image support is acceptable, things become far more complicated when considering the muscle images. This is due to the fact that for the left ventricle the separation of tissue and lumen is possible while the calf images do not exhibit clear separation between different muscles, and local deformations are far more pronounced in this structure.

4.1. Segmentation of the Left Ventricle

The automatic delineation of the LV is a critical component of computer-assisted cardiac diagnosis. Information with respect to the ejection fraction, the wall motion, and the valve behavior can be very useful toward predicting and avoiding myocardial infarction. The use of the shortest path algorithm along with shape matching was considered in [14], while active shape and appearance models were used for spatio-temporal heart segmentation in [20].

To evaluate the performance of the our representation, manifold construction and constrained search we consider 25 CT heart volumes, with an approximate voxel spacing of 1.5mm, for which 90 anatomical standard of reference landmarks, and a set of 1451 control points for the LV were available, besides the ground truth segmentation Fig.(4.a) from experts concerning the diastole as well as the systole.

4.2. Segmentation of Calf Muscle

The musculoskeletal modeling problem in medical imaging is not widely investigated in the literature, indeed few works have been dedicated to this issue [13, 2]. In fact the segmentation of individual muscles within a muscle compound depicted with *MRI* is an example that poses new challenges to automatic segmentation systems (Fig.(3)). Although dominated by the global anatomy, muscle deformation exhibits mostly locally consistent behavior, precluding the use of e.g., a global linear model. Muscle surfaces are only partially visible, with no prominent difference of tissue-properties between neighboring muscles. Border tissues in between muscles are only visible on specific loca-



Figure 3. Calf MRI data: (a) healthy (b) unhealthy case, (c) manual expert annotation.

tions, distributed in a very sparse and heterogeneous manner. State of the art medical segmentation methods mainly rely on a clearly defined topology, and an object boundary characterized by salient features (e.g., edges) [11]. Due to the topology free and compact modelisation property of our approach we can deal with the challenging calf muscles. In fact the border tissues in between muscles are only visible on specific locations, distributed in a very sparse and heterogeneous manner Fig.(3). Indeed muscle partially exhibit structures that can change dramatically between patients, or during the course of follow-up examinations. Nevertheless, and as highlighted by [2], musculoskeletal disabilities in general and Myopathies -as far as our work is concernedcould highly profit from this kind of studies and improve future treatments. Up to now the diffusion wavelet techniques has, so far, largely been confined to the signal processing field, and was not yet exploited for the need of medical image segmentation, to the best of our knowledge.

To assess our model, experiments were carried out over 25 MRI calf muscles divided into two groups: 20 healthy control patients and 5 unhealthy cases. For each volume there are 90 slices of 4mm thickness, and with voxel spacing 0.7812x0.7812x4 mm acquired with a 1.5T Siemens scanner. Standard of reference annotation by experts for the Medial Gastrocnemius (MG), was available (see Fig.(4.b)). Correspondences for 895 landmarks on the surfaces were obtained by an MDL based optimization [16].





(a) Ground truth segmentation of papillary muscles

(b) *T-1* MRI slice supervised segmentation of a human calf

Figure 4. Standard reference segmentation of respectively the left ventricle and the calf muscle.

4.3. Results: Model Reconstruction

During the evaluation of our framework, we model the shapes of the structures through diffusion wavelet representation in a leave-one-out cross validation strategy. To

	Heart Data	Calf Data
Gaussian model	1.6154	2.1277
DW Model with	0.0755	0.1485
spatial kernel		
DW Model with	0.1100	0.1796
Shape Map kernel		

Table 1. Full Landmark Reconstruction Error (in voxel) with regard to three different shape models for heart and calf data sets.



Figure 5. Reconstructed surfaces for Heart CT data using projected wavelet coefficients on the set of principal components that represent 99% of the total variance at level 1. The axial view surfaces represent the $\pm 3sqrt(\lambda_i)$ from left to right.

assess the diffusion modeling approach, we compute two measures: (i) reconstruction error, and (ii) search performance. We compare the reconstruction error for gaussian shape models, and the proposed diffusion wavelet modeling. We evaluate two different diffusion wavelet kernels: 1. the spacial proximity of landmarks, and 2. a kernel based on a shape map distance of the landmarks [17].

The main concern is to see how far our model is able to detect the local shape variations based on different kernels. To illustrate the orthomax representation, in the Fig.(5)we show the heart reconstructed surfaces using projected wavelet coefficients on the set of principal components at the first level, where the surfaces represent the coefficient values $\pm 3sqrt(\lambda_i)$ from left to right. Comparison of the reconstruction error between the diffusion wavelet model and the reference model is calculated as an average surface error for all test shapes. In Tab.(1), we report errors for a Gaussian reference model, and two diffusion wavelet models corresponding to different ways of deriving the diffusion operator. An optimal choice of the operator is subject of ongoing research. In Fig.(6) the reconstruction of the projected shape model (heart/muscle) is depicted, this reconstruction starts from the projection of the diffusion wavelet coefficients in respectively the first and the last level, and then extracting a new subset of coefficients from the eigenvectors that constitutes 99% of the variation in the correspondant level. To keep the variation in reasonable limits, the shape parameters are also restricted to $\pm 3\sqrt{\lambda_i}$.



Figure 6. Diffusion Wavelets Model Reconstruction. First row: Heart results and second: Calf muscle. Data, green: standard of reference segmentation, red: reconstruction result for a. fines scale and b. coarsest wavelet scale.

4.4. Results: Model Search

To assess the search behavior we compare our method with a standard gaussian shape model search in an active shape model search approach. We use an even sampling of the object surface, and gradients in the volume as texture description, and a sparse shape model as proposed in [12]. The latter uses a similar appearance model to the one used in this paper, and allows for the assessment of the effect of replacing the multivariate Gaussian landmark model, with the diffusion wavelet shape model. The error measure is the mean distance of the model landmarks between standard of reference and segmentation result Fig.(7). This gives also an indication of the displacement along the surface, which is relevant if the result is used for navigation. Models are initialized with minimal overlap to the target shape, and the accuracy of the final result was quantified by the mean landmark error between ground truth annotation and search result. For the quantitative comparison, results in Fig.(7) clearly show how the diffusion wavelet model outperforms the sparse model with standard parameterization for both anatomical data sets, with for example a mean value of 10.97 voxels for DW model over 13.72 error voxels for the sparse model in the calf data.

The diffusion wavelet model was able to recover the shape with superior accuracy. In the muscle data the standard search approach failed due to the ambiguous texture and local shape variability in large regions of the target shape. In Fig.(8) examples for standard, sparse model, and multi-scale diffusion wavelet based search are depicted. Note that one of the important points that distinguishes our



Figure 7. Boxplots of (a) Heart and (b) Calf Search Segmentation. Landmark Reconstruction Error (voxel) after finishing search phase, with comparison between three different search models; (1) our approach, (2) sparse model and (3) standard gaussian model.

methodology from robust ASMs, is that we learn the distribution of both image and shape information during the training phase to optimally exploit the anatomical properties of the data. This is not the case with for robust ASMs which for a given sampling consider a subset of the control points according to the observed image during search. In a typical segmentation scenario, our method runs approximately 56 seconds in average with non-optimized code implemented in Matlab 7.5, on a 2GHz DELL Duo Computer with 2Gb RAM.



Figure 8. Model search result for Heart muscle (upper row) and Calf muscles (down row). Data in green: standard of reference segmentation, in red: search results. For (a, d, g) standard gradient search approach, while (b, e, h) represent sparse shape models and finally (c, f, i) diffusion wavelet model.

5. Discussion

We present a multi-scale shape prior based on a diffusion wavelets shape representation and segmentation. Diffusion wavelet shape priors are able to take advantage of the subtle inter-dependencies in training data, by clustering coefficients based on correlation, and representing the topology of the structure by a diffusion kernel, instead of a fixed pre-defined manifold. We are using the orthomax criterion which is suitable for building sparse representations - particularly relevant in the case of the regions and pathologies studied in the paper - leading to localized modes of variation. The use of orthomax is motivated by obtaining an optimal subdivision of the shape parameterization.

We have shown that in the context of anatomical structures, the diffusion wavelet transformation is able to accurately and efficiently detect the locations and spatial scales of shape variations. The validation on detecting patterns on two complex medical data sets shows promising results indicating the advantage of using a learned model parameterization. We did also compare different search strategies on segmentation performance for both data sets.

Future work includes the integration of model learning approaches to learn the locations in non-annotated data in a weakly or unsupervised way. Furthermore a more in-depth study of optimal ways of deriving the diffusion operator during training will be explored. We would like to combine our representation/manifold with soft priors that penalize the distance from the manifold. Last the use of efficient optimization techniques in conjunction with the mentioned priors could lead to a flexible and powerful paradigm to represent and infer shapes of arbitrary topologies.

6. Acknowledgments

The authors would like to thank J. F. Deux, G. Basez, M. Maatouk and Pr. A. Rahmouni from Henry Mondor Universitary Hospital for providing clinical data.

References

- M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation. MIT Press*, 15:1373–1396, 2003.
- [2] S. Blemker, D. Asakawa, G. Gold, and S. Delp. Image-based musculoskeletal modeling: applications, advances, and future opportunities. *Journal of Magnetic Resonance Imaging*, 25(2):441–451, 2007.
- [3] G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Foundations of Computational Mathematics*, 5(1):1–58, 2005.
- [4] R. R. Coifman and S. Lafon. Diffusion maps. Appl. Comput. Harmon. Anal., 21:5–30, 2006.
- [5] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, F. Warner, and S. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. In *PNAS*, pages 7426–7431, 2005.

- [6] R. R. Coifman and M. Maggioni. Diffusion wavelets. Appl. Comput. Harmon. Anal., 21:53–94, 2006.
- [7] T. F. Cootes, G. Edwards, and C. Taylor. Active appearance models. In *ECCV*, pages 484–498, 1998.
- [8] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61:38–59, 1995.
- [9] D. Cremers, T. Kohlberger, and C. Schnörr. Nonlinear shape statistics in mumford-shah based segmentation. In *ECCV* (2), pages 93–108, 2002.
- [10] C. Davatzikos, X. Tao, and S. Dinggang. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22:414–423, 2003.
- [11] J. Duncan and N. Ayache. Medical image analysis: progress over two decades and the challenges ahead. *TPAMI*, 22:85– 106, 2000.
- [12] S. Essafi, G. Langs, and N. Paragios. Sparsity, redundancy and optimal image support towards knowledge-based segmentation. In *CVPR*, 2008.
- [13] B. Gilles, L. Moccozet, and N. Magnenat-Thalmann. Anatomical modelling of the musculoskeletal system from mri. *MICCAI*, pages 289–296, October 2006.
- [14] M. Jolly. Automatic segmentation of the left ventricle in cardiac MR and CT images. *IJCV*, 70(2):151–163, 2006.
- [15] H. Kaiser. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23:187–200, 1958.
- [16] G. Langs, R. Donner, P. Peloschek, and H. Bischof. Robust autonomous model learning from 2d and 3d data sets. In *Proc. of MICCAI'07*, 2007.
- [17] G. Langs and N. Paragios. Modeling the structure of multivariate manifolds: Shape maps. In CVPR, 2008.
- [18] B. Luo and E. R. Hancock. Iterative procrustes alignment with the em algorithm. *Image and Vision Computing*, 2002.
- [19] Y. Meyer. Wavelets Algorithms and applications. Applied Mathematics, 1993.
- [20] S. Mitchell, J. Bosch, B. Lelieveldt, R. Van der Geest, J. Reiber, and M. Sonka. 3-d active appearance models: segmentation of cardiac mr and ultrasound images. *IEEE Trans. Medical Imaging*, 21(9):1167–1178, 2002.
- [21] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum. Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Trans Med Imaging*, 26(4):598–618, 2007.
- [22] M. Rousson and D. Cremers. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In *MICCAI*, volume 1, pages 757–764, 2005.
- [23] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.
- [24] M. Stegmann, K. Sjostrand, and R. Larsen. Sparse modeling of landmark and texture variability using the orthomax criterion. In SPIE Medical Imaging, volume 6144, 2006.
- [25] M. Taron, N. Paragios, and M.-P. Jolly. Registration with uncertainties and statistical modeling of shapes with variable metric kernels. *IEEE TPAMI*, 31(1):99–113, 2009.
- [26] K. Zhu, Y. Wong, W. Lu, and J. Fuh. A diffusion wavelet approach for 3-d model matching. *Computer-Aided Design*, 41(1):28–36, 2009.