

Left Ventricle Segmentation Using Diffusion Wavelets and Boosting

Salma Essafi^{1,2}, Georg Langs³, and Nikos Paragios^{1,2}

¹ Laboratoire de Mathématiques Appliquées aux Systèmes,
Ecole Centrale de Paris, France

² GALEN Group, INRIA Saclay-Ile de France, Orsay, France

³ Computational Image Analysis and Radiology Lab, Department of Radiology,
Medical University of Vienna, Austria

salma.essafi@ecp.fr, georg.langs@meduniwien.ac.at,
nikos.paragios@ecp.fr

Abstract. We propose a method for the segmentation of medical images based on a novel parameterization of prior shape knowledge and a search scheme based on classifying local appearance. The method uses diffusion wavelets to capture arbitrary and continuous interdependencies in the training data and uses them for an efficient shape model. The lack of classic visual consistency in complex medical imaging data, is tackled by a manifold learning approach handling optimal high-dimensional local features by Gentle Boosting. Appearance saliency is encoded in the model and segmentation is performed through the extraction and classification of the corresponding features in a new data set, as well as a diffusion wavelet based shape model constraint. Our framework 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. Promising results are reported for heart CT data sets, proving the impact of the soft parameterization, and the efficiency of our approach.

1 Introduction

Data acquired by medical imaging modalities has a level of richness that needs computer based methods to extract relevant information in a consistent and efficient manner. The automatic and accurate delineation of the left ventricle (LV) is a prominent example for 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. Existing segmentation approaches include the use of a shortest path algorithm along with shape matching which was considered in [1], or an alternative shape representation using level set functions was proposed in [2].

These methods depend heavily on the accuracy of the inter-subject registration for group comparison and the parameterization of the shape. A promising line of research considering wavelets for the representation of shapes was initiated in [3] by build hierarchical active shape models of 2-D anatomical objects using 1-D wavelets, which are then used for shape based image segmentation. A further extension was proposed in [4]

where spherical wavelets are used to characterize shape variation in a local fashion in the space and frequency domain.

Two crucial components of image model based methods are the parameterization of the shape manifold, and the capturing and representation of the appearance in the training and search data. The model based segmentation approach proposed in this paper accounts for the systematic behavior of shape variation and image support in anatomical structures, with a parameterization that goes beyond pre-defined reference manifolds. For the parameterization of complex structures it is worthwhile to not rely on a reference manifold with an a priori topology, but to learn the appropriate topology from the training data. For this we have to determine the intrinsic topology of a shape for which multiple examples are available, and have to encode this information in the shape model, to use it in the representation and during segmentation.

We propose a method that integrates local voxel classification and global search models. We model and parameterize shape variation of structures with arbitrary topology, by using diffusion wavelet shape models [5] to represent the shape variation with a learnt parameterization based on mutual distance. The approach deals with complex and soft connectivity properties of objects by encoding their interdependencies with a diffusion kernel [6]. The topology is learned from the training data instead of using a priori choices like e.g., a sphere and represents the shape variation by means of diffusion wavelets [7]. A detailed explanation of diffusion wavelet shape models, including variants of the parameterization can be found in [8].

During search this model is used together with a GentleBoost classifier [9] trained on the local appearance of the individual landmarks describing the anatomical structure. The method obtains an accurate delineation of partially visible surfaces and complex texture, that cannot be achieved with registration based methods. The shape representation is based on a finite set of landmarks, that can be repeatedly identified and exhibit significant differentiation to the background on different examples of the anatomical structure, and more particularly for CT cardiac volumes. During the search the hierarchical diffusion wavelet shape model [8] is fitted to new data based on local appearance captured by the classifier. Related approaches combining local features with standard shape models are [10], or [11]. The method computes a local feature vector for every voxel and maps it via a GentleBoost classifier [9] to a probability that the voxel belongs to a specific landmark in the object. The classifier is trained from the training data set segmentations. The probabilistic output is constrained by the shape model. The mapping onto the diffusion wavelet coefficient space ensures valid results with regard to the training data. The result of this procedure is a probability for each voxel regarding its match to the structure to be segmented, conditioned on both local and global information. We report results on CT left heart ventricle data sets, that illustrate the impact of the soft parameterization, as well as the global classifier based search.

2 Hierarchical Shape Model Building

We model the shape variation observed in the training data by means of diffusion wavelets. Wavelets represent a robust mathematical tool for hierarchically decomposing functions into different frequency components. We refer the reader to [12] for complete

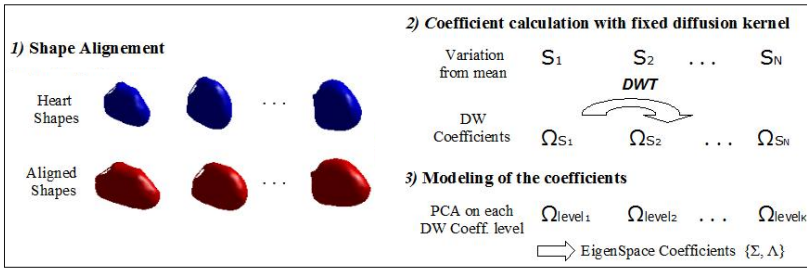


Fig. 1. Scheme of Diffusion Wavelet Coefficient Process

description of the theory. The major advantage of wavelets is the compact support of basis functions on one hand, and on the inherently hierarchical representation based on multi-resolution modeling of processes at different spatial and temporal scales. The diffusion wavelet technique introduced by [7] associates multi-scale representation of training data as well as the topological information captured by means of a diffusion kernel [6]. Diffusion wavelets enable local as well as global variation detection, which makes it useful and suitable for our application.

For the modeling of the variation, we consider the heart volumes as a finite set of landmarks. Starting from m landmark positions, $\mathbf{H}_i = \langle \mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_m^i \rangle$, are known in n training volume images $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N$. Our data comprises $\mathcal{H} = \{\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_N\}$, where $\mathbf{x}_j^i \in \mathbb{R}^d$, and we refer to $\mathbf{H}_i \in \mathbb{R}^{dm}$ a shape. Since we are only interested in the non-rigid deformation, all anatomical shapes are aligned by Procrustes analysis, which produces the series of examples H_i^p , from which we compute the mean shape \bar{H}^p (Fig. 1). Once registered, the shapes are used through their deviation S_i from the mean shape, $S_i = H_i^p - \bar{H}^p$, where \bar{H}^p represents the volumes mean shape.

We now specify a topology over the set of landmarks. For this we use a hierarchical geometric graph framework introduced in [6]. It applies the concept of diffusion to capture mutual relations between nodes in a Markov chain, and derives the global structure of a shape. In our case, this structure is the neighborhood relation between landmarks of the shape, that determines the domain upon which the wavelet representation is built. Diffusion maps grant a canonical representation of high-dimensional data. By this we are able to represent spatial relations as well as the data behavior. The structure is encoded in a diffusion operator $\Delta \in \mathbb{R}^{m \times m}$. Combining those two diffusion approaches leads us to a prior knowledge of global and local training population variation [8]. The diffusion operator Δ is built on the set of points embedded in a metric space utilizing their mutual distance in the mean shape, and reflects all pairwise relations between individual points in the shape set.

After defining the diffusion operator Δ , we build the according *diffusion wavelet tree*. For this we apply general multi resolution construction for efficiently computing, representing and compressing Δ^{2^j} , for $j > 0$. The latter are dyadic powers of Δ , and we use them as dilation operators to move from one level to the next, which is a simple way of compressing high orders of the diffusion operator. The process of constructing the diffusion wavelet, the tree and the coefficients is described in detail in [7].

Once the tree is built based on the diffusion operator, we can compute the diffusion wavelet coefficients Ω for each shape S_i , so that $\Omega_{S_i} = \Psi^{-1}S_i$, where Ψ represents the diffusion wavelet tree. Hence we can rebuild our shape as $H_i^p = \bar{H}^p + \Psi\Omega_{S_i}$.

Now we move to a representation scale by scale, in order to construct a model of the variation at each level for all the population training. We gather the low frequency information in the coarser level, while localized variations will be detected through high level coefficients in the multi scale representation. We define $\Omega_{level\ j}$ at every scale j , with $(1 \leq j \leq K)$, such as $\Omega_{level\ j} = \{\Omega_{S_i/level=j}\}_{i=1\dots N}$. Afterward we perform principle component analysis (PCA) for the coefficients of all scales.

The eigenspace resulting from this PCA will be referred as $\{\Lambda, \Sigma\}$, where more precisely will have $\Sigma = \{\sigma_j\}_{j=1\dots K}$, and the corresponding eigenvalues $\Lambda = \{\lambda_j\}_{j=1\dots K}$ of the covariance matrix of the diffusion wavelets coefficients at each level j , and the according coefficients $\Omega_{level\ j}^*$ that represent each training shape in this coordinate system. Hence in each level the coefficients would be expressed such as:

$$\Omega_{level\ j} = \bar{\Omega}_{level\ j} + \sigma_j (\sigma'_j \cdot \Omega_{level\ j}^*) \tag{1}$$

Finally we can activate the shape reconstruction process: (1) we compute the diffusion wavelet coefficients $\Omega_{S_i\ Rec}$ in each level, (2) we remodel the shape based on the diffusion wavelet tree, and get the reconstructed shape: $\Upsilon_i^p = \bar{H}^p + \Psi\Omega_{S_i\ Rec}$.

3 Segmentation Based on Image Information and the Model Prior

The segmentation of the LV is challenging mostly due to the similar visual properties with the other chambers of the heart cavity, as well as the presence of papillary muscles. The use of edge-driven terms with regional statistics along either with deformable contours or active shape and appearance models. In the first case, computational complexity is an issue and the proper handling of papillary muscles is problematic. In the second case, one has to deal with either the linearity of the sub-space or the fact that building appearance modes requires appearance normalization and too many samples. We adopt recent developments in machine learning that explores the use of weak classifiers and arbitrary image features. In the context of the heart muscle, our feature space involves the (i) gradient phase and magnitude, (ii) structure tensor plus their (iii) curvature [13] and (v) the responses to Gabor filters with different phases and orientations.

3.1 Learning a Classifier for Appearance Modeling

Starting from this feature space, we apply Gentle Adaboost [9] to obtain a local appearance prior for the search in new data. The boosting process aims to build a strong classifier by combining a number of weak classifiers, which need only be better than chance. For this we call upon a sequential learning process: at each iteration, we add a weak classifier. It is the basic learning algorithm introduced by Viola, Jones in [14]. Our classification problem evolves as a two class training set that can be represented as: $S = \{(x_i, y_i)\}_{i=1}^l \subset \mathbb{R}^N \times \{-1, +1\}$.

Given these classifiers, we use them to locate landmarks during the segmentation process. The classifiers detect landmarks present on the ventricle muscle wall against

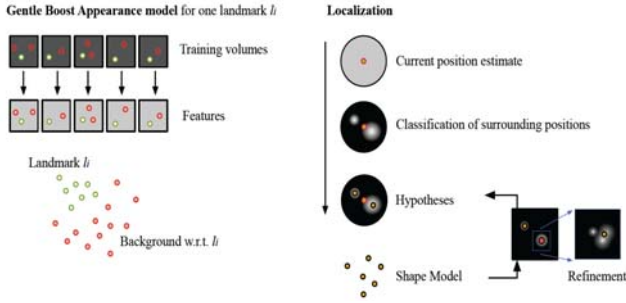


Fig. 2. Scheme of the appearance model: Based on local features, and a classifier we can assign each position in the volume an evidence value for landmarks presence. This results in a set of hypotheses for landmark positions, that are verified by the shape model constraint.

background. This is a very different strategy in comparison to standard search methods [15]. The main search strategy is: extract features from the volume, for each landmark obtain a few candidate positions with a very strong classifier response, fit the DW model to these candidates, and determine the candidate configuration with the highest plausibility with regard to the shape prior. After continue with the local search at the current landmark estimates constraint by the DW model. During the shape model fitting we check which candidates have the highest plausibility with the trained DW model.

In Fig. 2 the scheme of the model search is depicted. For each landmark the search volume \mathbf{V} is projected into a hypothesis space \mathbf{V}_i^H that reflects the evidence for the landmark presence for each point in the volume. This results in a position hypothesis $\hat{\mathbf{x}}_i$ for each landmark. The set of landmark hypotheses $\langle \hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_n \rangle$ is tested with the diffusion wavelet shape model, resulting in a position prediction for each landmark. These predictions are used to generate new hypotheses based on the local image support $\mathbf{V}_i^{H'}$ and the shape model. The hypothesis space is the classifier response on each position in the volume. During the progressing search we just consider the neighborhood of the current landmark location estimate during the last iteration.

3.2 The Segmentation Algorithm

Let us summarize the learning and search concepts introduced in this paper. The method consists of a training phase and an execution step. First, the shape model and parameterization, and the local classifiers for the appearance representation are learnt. During search they are used to locate and segment structures in new image data.

Learning: During the training both geometry and appearance of the structure of interest are learned.

- Given n examples of the structure of interest location and the corresponding images, we represent the shape variability through diffusion wavelets.
- Using the same examples, we compute the selected feature images for each training example for different resolutions. For each landmark, at each resolution, we construct a set of training samples containing local features and

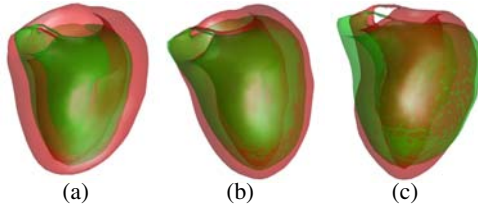


Fig. 3. Reconstructed surfaces for Heart CT data using projected wavelet coefficients on PCA eigenspace, representing 99% of the total variance at finest level. The surfaces are showing the $\pm 3\text{sqrt}(\lambda_i)$ from left to right. Figure is dedicated to the sagittal view.

corresponding labels which indicate if the position is the landmark location or the background, i.e., whether it is on the ventricle wall or on the background. Let's note here that background voxels are chosen randomly in the volume except the particular landmark positions for training.

- To train for the fine local differentiation we take into consideration only the neighborhood of every landmark candidate in each training image. We train a classifier for each landmark and retain only the ones with solid performance or wide-margins between the different classes.

Segmentation: Using both geometric and appearance priors and the corresponding feature space in the image we perform the structure delineation as follows: the process is initiated with the mean shape, and proceeds in an iterative manner,

- Perform a local search for the most probable landmark positions using the trained classifiers
- Constrain the solution using the diffusion wavelet coefficient constraints, and repeats the previous search steps until convergence.

This results in landmark location estimates in the search image, that are based on the appearance, and the shape constrained learned during the training phase.

4 Experimental Validation

To assess the performance of our approach, we consider a data set that includes 25 CT volumes of the heart, with an approximate voxel spacing of 1.5 mm, for which 90 anatomical standard of reference landmarks, and a set of 1451 control points for the left ventricle was available, in addition to the ground truth segmentation from experts concerning the diastole as well as the systole.

We have run our algorithm in a leave-one-out cross validation fashion. For the diffusion wavelet building part, we obtain 9 diffusion wavelet levels of decomposition for the shape prior. As for the initialization of our framework, we used the mean shape displaced by a random translation of 30 mm.

To evaluate the efficiency of our method, we computed two error measures: (i) the Hausdorff distance revealing the maximum error between the standard of reference and our model reconstruction, as well as (ii) mean distance error of the detected landmarks. In Fig.4.a, one can see that the Hausdorff distance error decreases with an increasing

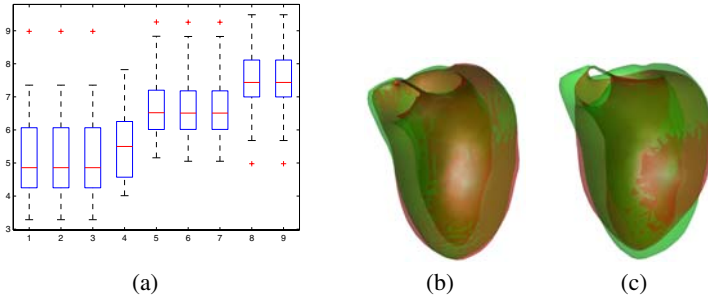


Fig. 4. Multiscale Diffusion Wavelets Reconstruction. (a) Hausdorff Error Distance (in voxel) of reconstructed heart at each diffusion scale for all data in the training set. (b) Data, green: ground truth segmentation, red: reconstruction result for finest scale and (c) coarsest wavelet scale.

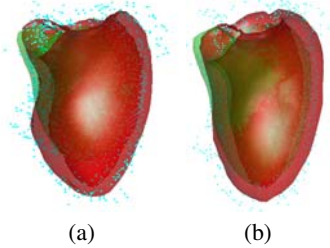


Fig. 5. Model search result for Heart muscle. Ground truth in green, in red: search results. (a) standard Gaussian search approach, and (b) method presented in this paper.

number of diffusion wavelet levels used for reconstruction. When we consider the mean reconstruction error over all data, we reach a distance of 2.2313 voxel in the image for the finest level, while as for the coarsest level we obtain 2.7073 voxel. The comparison of detection results for different numbers of levels used during reconstruction can be seen in Fig.4. Note that diffusion wavelets have been shown to outperform standard Gaussian models in terms of search error in [5] on muscle MRI data.

During the search validation experiments, we consider a multi-resolution approach for each landmark patch which goes from 5×5 pixels to 20×20 pixels in 4 steps. We obtain 200 landmarks candidates, for 15 training hearts and 10 testing examples. Experiments were carried out using Gentle Adaboost, which is adequate to deal with a large number of negative examples as well as the rather limited size of our training set.

In the quantitative assessment of the search/segmentation algorithm explained in Sec.3, we obtain a lowest error of 4.72 voxel between ground truth and relative segmented volume. In a typical segmentation scenario, the method takes approximately 68 seconds in average through non-optimized code implemented in Matlab 7.5, on a 2GHz DELL Duo Computer with 2Gb RAM. One should note here that we are working toward search in very large data sets, while searching for small complex structures, thus the efficiency of gradient descent of ASM is limited. In an ideal case one would combine the trade off between the reconstruction accuracy and the classification error to choose the best candidate for the search segmentation.

5 Conclusion

In this paper we propose a 3D hierarchical shape prior segmentation framework based on diffusion wavelets and local appearance classifiers. The diffusion wavelets are able to represent subtle inter-dependencies in the training data, by clustering coefficients, and representing the topology of the structure by a diffusion kernel, instead of a fixed pre-defined manifold. The conjunction of the diffusion wavelet constraint with a search method based on a GentleBoost classifier leads to an effective segmentation scheme. It can deal with ambiguous appearance and complex structures. Future work will focus on extensive evaluation, and the integration of efficient optimization techniques in addition to the studied priors to obtain a more flexible and powerful paradigm for representing shapes of arbitrary topologies, and the search in large data sets.

References

1. Jolly, M.: Automatic segmentation of the left ventricle in cardiac MR and CT images. *IJCV* 70(2), 151–163 (2006)
2. Paragios, N.: A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis. *International Journal of Computer Vision* 50(3), 345–362 (2002)
3. Davatzikos, C., Tao, X., Dinggang, S.: Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging* 22, 414–423 (2003)
4. Nain, D., Haker, S., Bobick, A., Tannenbaum, A.: Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Trans. Med. Imaging* 26(4), 598–618 (2007)
5. Essafi, S., Langs, G., Deux, J.F., Rahmouni, A., Bassez, G., Paragios, N.: Wavelet-driven knowledge-based MRI calf segmentation. In: *Proceedings of ISBI* (2009)
6. Coifman, R.R., Lafon, S.: Diffusion maps. *Appl. Comput. Harmon. Anal.* 21, 5–30 (2006)
7. Coifman, R.R., Maggioni, M.: Diffusion wavelets. *Appl. Comput. Harmon. Anal.* 21, 53–94 (2006)
8. Essafi, S., Langs, G., Paragios, N.: Hierarchical 3d diffusion wavelets shape priors. In: *IEEE International Conference in Computer Vision, ICCV 2009* (2009)
9. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: A statistical view of boosting. *Annals of statistics*, 337–374 (2000)
10. Scott, I.M., Cootes, T.F., Taylor, C.J.: Improving appearance model matching using local image structure. In: Taylor, C.J., Noble, J.A. (eds.) *IPMI 2003. LNCS*, vol. 2732, pp. 258–269. Springer, Heidelberg (2003)
11. Qian, Z., Metaxas, D.N., Axel, L.: A learning framework for the automatic and accurate segmentation of cardiac tagged MRI images. In: *CVBIA*, pp. 93–102 (2005)
12. Meyer, Y.: *Wavelets - Algorithms and applications*. Applied Mathematics (1993)
13. Rieger, B., Timmermans, F.J., van Vliet, L.J., Verbeek, P.W.: On curvature estimation of iso surfaces in 3d gray-value images and the computation of shape descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(8), 1088–1094 (2004)
14. Viola, P., Jones, M.: Robust real-time object detection. *International Journal of Computer Vision* (2001)
15. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. *Computer Vision and Image Understanding* 61, 38–59 (1995)