Sparsity, Redundancy and Optimal Image Support towards Knowledge-based Segmentation

Salma Essafi Georg Langs Nikos Paragios
Laboratoire MAS, Ecole Centrale Paris, Châtenay-Malabry, France
Equipe GALEN, INRIA Saclay - Île-de-France, Orsay, France
salma.essafi@ecp.fr, georg.langs@ecp.fr, nikos.paragios@ecp.fr *

Abstract

In this paper, we propose a novel approach to model shape variations. It encodes sparsity, exploits geometric redundancy, and accounts for the different degrees of local variation and image support. In this context we consider a control-point based shape representation. Their sparse distribution is derived based on a shape model metric learned from the training data, and the ambiguity of local appearance with regard to segmentation changes. The resulting sparse model of the object improves reconstruction and search behavior, in particular for data that exhibit a heterogeneous distribution of image information and shape complexity. Furthermore, it goes beyond conventional image-based segmentation approaches since it is able to identify reliable image structures which are then encoded within the model and used to determine the optimal segmentation map. We report promising experimental results comparing our approach with standard models on MRI data of calf muscles - an application where traditional image-based methods fail - and CT data of the left heart ventricle.

1. Introduction

Segmentation is a fundamental problem in image processing, medical image analysis and computer vision. In the most general case one would like to create a partition of the image in regions with similar features. However, such a problem is ill-posed mostly because observations are defined at the pixel level, while the conceptual grouping - or segmentation - refers to the notion of an entire object. Knowledge-based methods tackle this demand by first determining some prior density on the space of solutions and then constraining the solution in a new image to this density. Such a process involves three key aspects: (i) shape representation, (ii) modeling of shape variations and (iii) inference from new data.

The vast majority of existing approaches are explicit (landmark-based), and deduct critical points along with an interpolation strategy towards describing the entire shape. Examples are active shape models [6], spline-driven representations [2], the triangulation of surfaces [11], or wavelet-based representations [16], to name a few. The selection of the landmarks as well as the interpolation strategy is an important challenge towards recovering the least possible complex representation with the best possible geometric reconstruction of the object under consideration.

Once the representation has been been defined, the next step consists of learning a manifold on the resulting space from a set of training examples. Linear subspaces, parametric as well as non-parametric densities have been considered to model shape variation through the observed global distribution of the landmarks within the training examples. Multi-variate Gaussians, as well as kernel-based representations of fixed and variable bandwidth are examples for the representation of the density.

During search, the inference step consists of recovering an instance of the learnt representation which is part of the model manifold and is best supported by the observed image features. Boundary-based methods are seeking the optimal instance of the model guided by image-discontinuities, region-driven approaches aim to separate the global intensity characteristics of the object from the background, while feature-driven methods seek to learn patterns of support in the image. One can refer to a number of surveys published for deformable models [15], markov random field-driven [21], active appearance models [5] and minimal paths and fast marching techniques [20]. Often, these three steps are treated independently. Once the representation has been determined, an assumption on the statistical model is made and the parameters of the manifold are determined from the training set. This manifold is then used along with the image features for object extraction. One should make the following observations: a strong dependency exists (i) be-

*This work was partially supported by Association Française contre les Myopathies (AFM: http://www.afm-france.org) under the DTI-MUSCLE project, and by the Region Ile-de-France.

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

tween the representation and the model since by changing the distribution of control points, the model can be very different with regard to its representative capabilities; (ii) between the representation and the segmentation since image features are often computed based on the representation, and (iii) between the model and the segmentation (for example manifold-enhanced vs manifold-enforced methods).

The aim of this paper is to propose a novel integrated approach that addresses the above-mentioned interdependencies between tasks, and uses them to improve the model representation. We propose a unified modeling approach for both shape and appearance in computer vision as well as in medical imaging.

The method obtains a sparse model of objects or anatomical structures that takes the local statistical modeling and appearance behavior in the training set into account. We compare the reconstruction and search behavior with standard shape models, that neglect these properties, and represent objects, regardless of the reliability and complexity of shape and texture behavior in the training population. The model based segmentation is based on a sparse set of landmarks that can be uniquely identified in new data during search. The model is built based on a set of training examples for which expert annotations are available. It integrates knowledge about local appearance, shape variability and the ambiguity of image data to achieve a segmentation performance equal to or superior to a medical expert.

The remainder of the paper is organized as follows: in Sec. 2 the optimal shape representation and model construction are presented. Sec. 3 deals first with the inference and the optimal use of image support, while afterward we introduce the application context, that is the segmentation, and report experimental results. Finally Sec. 4 concludes the paper with discussion and different openings.

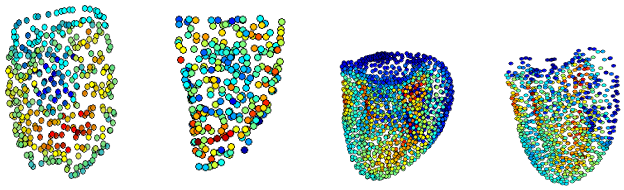
2. Sparse Shape Models

Sparse shape models learn a representation and a corresponding reconstruction mechanism from a set of training examples. The sparse model is built based on the statistical behavior of the training shapes and the distribution of appearance information in the training data. In the following we will formulate the framework, and explain how to derive an optimal sparse representation from training examples. Subsequently the method for the reconstruction of the entire modeled structure from the sparse representation, and the search procedure will be explained.

Let us consider without loss of generality a shape representation that consists of a finite set of landmarks. Given a set of n training volumes and their corresponding segmented structure

$$\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_n. \tag{1}$$

our knowledge about the data comprises m landmark positions in each of the examples, that are located at consistent positions with regard to the object or structure of interest. The number of landmarks can be high, up to a dense sampling of shape surfaces. Landmarks are not constraint to anatomically salient points, but can be distributed on manually segmented training examples by methods like those proposed in [7, 12]. The landmark positions can be found using a number of approaches, like for example the minimum description length criterion. Landmarks do not have to be located on a single surface or manifold, but can define arbitrary structures and deformation fields [19]. For each example $\mathbf{I}_i$, the landmarks are located at the positions



(a) (b) (c) (d)

Figure 1. Sparse shape models: Calf muscle: (a) colour coded density in the shape diffusion map according sparse landmark distribution (b). Left ventricle: (c) shape diffusion map density and (d) sparse sub-sampling.

$$\mathbf{V}_i = \langle \mathbf{x}_1^i, \mathbf{x}_2^i, \ldots \mathbf{x}_m^i \rangle. \tag{2}$$

where $\mathbf{x}_i \in \mathbb{R}^d$. We call $\mathbf{V}_i \in \mathbb{R}^{dm}$ a shape, and denote the set of shapes in the training set by

$$\mathcal{V} = \{\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_n\}. \tag{3}$$

This data defines a shape manifold, that can be associated with geometric and image support. The local image information and manifold geometry can be used to reduce the number of landmarks. This should happen while satisfying two conditions: (i) the preservation of the maximum amount of information about the shapes in the training set, i.e., the reconstruction ability, and (ii) the optimal use of the image information during inference in new data. This is equivalent with recovering a reduced rank representation

$$\hat{\mathcal{V}} = \{\hat{\mathbf{V}}_1, \hat{\mathbf{V}}_2, \ldots, \hat{\mathbf{V}}_n\}. \tag{4}$$

where $\hat{\mathbf{V}}_i \in \mathbb{R}^{d(m')}$, and $m' \ll m$. $\hat{\mathbf{V}}_i$ consists of a landmarks subset defining the shape, and a corresponding reconstruction function $P : \mathbb{R}^{d(m')} \rightarrow \mathbb{R}^{dm}$, $\hat{\mathbf{V}}_i \mapsto \mathbf{V}_1 + \mathcal{R}$ where $\mathcal{R}$ is a residual error, that should be minimal.

We will first discuss how to obtain this representation $\hat{\mathcal{V}}$ based on the shape and appearance behavior in the training set, in order to obtain optimal search ability. we consider here a multivariate Gaussian shape model as used in [7].

2.1. Shape Maps and Redundancy

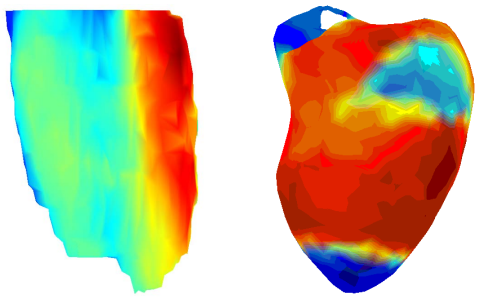
We view finding an optimal shape representation as an optimal sampling with respect to the variations being ob-

served in the training data. It should have low density in regions that behave in a redundant manner, and high density in regions that exhibit uncorrelated deformation behavior in the training set. Analogously to a uniformly distributed sampling in real space, which covers the object evenly, the sparse representation has to cover the object evenly with regard to the *information* contained in each sampling point. To achieve this, we have to capture the coherence of the behavior of shape regions in the training examples. In [13] shape maps are introduced. They provide for a shape population metric, that captures the interdependencies in the behavior of landmarks. We use the concept of shape maps to derive an optimal sampling. Let us consider a Markov chain consisting of m nodes, each corresponding to one landmark, and edges with a value $p_k(i,j)$ between the nodes that correspond to the minimal description lengths [17] of models encompassing the two landmarks i and j and $k-2$ other landmarks. The description length L is the number of bits, that are necessary to communicate a model $\mathcal{M}$, the data D (in our case landmark positions) encoded with help of this model, and a residual error: $L(D,\mathcal{M}) = L(\mathcal{M}) + L(D|\mathcal{M}) + \mathcal{R}$. The data term is associated with the reconstruction error, while the model term penalizes over-fitting through the use of expensive (in terms of number of parameters) models. In our case it provides information about the compactness of models describing the joint variation of the landmarks i and j, or equivalently about the redundancy in their position information in the training set. We expect low values for edges between landmarks, that behave in a coherent way. If $d_k(i,j)$ is the minimal description length [7], then the normalized graph Laplacian construction [3] allows us to construct a reversible Markov chain from the symmetric graph defined by the nodes and edges $p(i,j) = k(i,j)/\sum_j k(i,j)$, where $k(i,j) = e^{-\frac{d_k(i,j)}{\epsilon}}$. This Markov chain is given by the non-symmetric matrix P with entries $p(i,j)$, or its powers P^t that correspond to an increasing time in the chain, and to the according propagation of probabilities. The Markov chain captures the shape variation behavior by connecting groups of coherent landmarks with high-valued edges, while having low value edges between landmarks, that share only limited mutual information. The pairwise relation between landmarks is captured by the according *diffusion distance*. An eigenvalue analysis of P allows to generate a *diffusion map* [4] - or *shape map* - a metric space, in which a *diffusion distance* parameterized by t

$$D_t(i,j) = \sum_u \frac{(p_t(i,l) - p_t(j,l))^2}{\pi(l)}, \quad \pi(i) = \frac{d(i)}{\sum_j d(j)}, \tag{5}$$

becomes the Euclidean distance between the images of i and j, $\mathbf{\Psi}_t(i)$ and $\mathbf{\Psi}_t(j)$,

$$\|\mathbf{\Psi}_t(i) - \mathbf{\Psi}_t(j)\| = D_t(i,j). \tag{6}$$



(a) (b)

Figure 2. Image support: a. muscle surface, b. left ventricle.

Once the shape map is generated, the density estimation can be performed using an Euclidean approach. The i landmark has an image in the map, which we denote by $\Psi_i = \mathbf{\Psi}_t(i)$, while t is used to describe the entire space. In such a context, the diffusion map $\mathcal{S}$, is a metric space and therefore we can estimate the density d_{Ψ_i} of the landmark images $\Psi_i \in \mathcal{S}$ for each point. The density relates to the number of landmarks, that can be encoded by the same model while retaining low description length. It is a measure of redundancy [22], since a model that represents the shape variation of a set of landmarks with images Ψ_i in a small neighborhood in $\mathcal{S}$ is compact - according to the generation of the Markov chain - and indicates that the mutual information that landmarks carry about each other is high. In Fig. 1 shape diffusion maps, and densities are depicted for a set of calf muscles, and a set of left ventricles. A detailed derivation of shape maps is given in [13]. We aim for a sampling, in which each of the landmarks i carries an equal amount of mutual information about about other landmarks k, which have diffusion map images Ψ_k in its neighborhood. This would result in a uniform distribution of images Ψ_i in $\mathcal{S}$.

To conclude: the shape map assigns each landmark a position Ψ_i. The distance between Ψ_i and Ψ_j in the shape map corresponds to the coherence of the behavior of landmarks in the training set. We aim at a uniform sampling in the shape map, so that the mutual information landmarks carry about each other is evenly distributed. The next step is to add appearance information to the map.

2.2. Image support

The shape map represents the shape variation structure of the training examples. The appearance information, that is used during search, is not distributed evenly on the entire object, too. In the case of muscle MRIs only a small ratio of the surface carries distinctive appearance (Fig. 3), that allows for a separation between background and foreground. To account for this variability we calculate the *image support* at each landmark position during training. If we can assign a value to a landmark relating to the distinc-

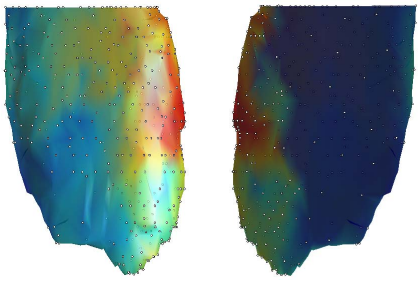


Figure 3. Surface of a calf muscle: image support on the outer and inner part, and the sparse model points.

tiveness of the local texture in the training set $\mathcal{V}$ we can further differentiate the representation $\hat{\mathcal{V}}$. Conceptually, the model should use landmarks with salient appearance for the inference from the data, while reconstructing the remaining parts of the shape according to a reconstruction mechanism based on the shape prior.

For each landmark i we denote by g_i the image support in the training set. g relates to the chosen search strategy. Since we employ local texture patches, we derive g based on the distinctiveness of the texture at the landmark position. In Fig. 2 the image support for calf muscles, and left ventricles is depicted. We calculate the correlation of texture appearance in the vicinity of landmark positions in the training set. For distinctive features, the correlation can be expected to show a peak at the correct position. Let $b(\mathbf{x}_i^j)$ be the learned texture patch at the correct landmark position $\mathbf{x}_i$ in the training example $\mathbf{I}_j$, and for landmark positions in a local neighborhood $\mathcal{N}$ let $Q_i^j(\mathbf{x})$ be the correlation between the patch $b(\mathbf{x})$ and $b(\mathbf{x}_i^j)$ normalized within the neighborhood, i.e. $\int_{\mathbf{x}\in\mathcal{N}} Q_i^j(\mathbf{x}) = 1$, then the image support is

$$g_i = \sum_{j=1}^{n} \left(\frac{Q_i^j(\mathbf{x})}{\int_{\mathbf{x}\in\mathcal{N}\setminus\mathbf{x}_i^j} Q_i^j(\mathbf{x})} \right). \tag{7}$$

That is, for a landmark in $\mathcal{V}$ the image support is calculated from the local appearance behavior at the corresponding positions in the training set.

2.3. Sparse sampling of the data

Given a metric space $\mathcal{S}$ that captures the statistical shape behavior, an according density d_{Ψ_i} and an image support g_i for each landmark we obtain a sparse sampling by minimizing the integral of absolute gradient in the map $\mathcal{S}$,

$$\mathcal{C}(\hat{\mathcal{V}}) = \int_{\mathcal{S}, i\in\hat{\mathcal{V}}} |\nabla(d_{\Psi_i}/\gamma g_i)|, \tag{8}$$

by choosing a subset of landmarks. That is, the function reaches a minimum if an even distribution of landmark images weighted by g_i is obtained. This distribution favors landmarks that have high image support in the training set, while integrating the statistical shape modeling and reconstruction properties of individual landmarks.

Given a diffusion map $\mathcal{S}$, a set of object landmarks images $\mathcal{Y}_0 = \{\Psi_1, \dots \Psi_m\} \subset \mathcal{S}$, $\hat{\mathcal{Y}}_0 = \emptyset$, and the according densities $d_{\Psi_1}, \dots d_{\Psi_m}$, and a value r, we perform the sparse sampling in the following iterative way: 1. choose $i : d_{\Psi_i} = \max(\{d_{\Psi_j} : \Psi_j \in \mathcal{Y}_t\})$; 2. set $\mathcal{Y}_{t+1} = \mathcal{Y}_t \setminus \{\Psi_i \cup \Psi_j : \|\Psi_j - \Psi_i\| \leq r/\gamma g_i\}$, $\hat{\mathcal{Y}}_{t+1} = \hat{\mathcal{Y}}_t \cup \Psi_i$, and iterate until $\mathcal{Y} = \emptyset$. The value r controls the mean density of the sparse representation.

This results in a set $\hat{\mathcal{Y}}$ and a corresponding set of landmarks $\hat{\mathcal{V}}$ that forms the sparse model representation, in which the mutual information between landmarks and the appearance information at landmark positions is distributed evenly. In the following we will explain how to reconstruct the entire object $\mathbf{X}$ from the sparse representation $\hat{\mathbf{X}}$ using the diffusion map $\mathcal{S}$. In Fig. 3 a sparse sampling for calf muscles is depicted together with the color coded image support.

2.4. Reconstruction

The reconstruction of the shape consists of inferring the positions of the entire shape $\mathbf{V_i} = \langle \mathbf{x}_1^i, \mathbf{x}_2^i, \dots \mathbf{x}_m^i, \rangle$ from the sparse representation $\hat{\mathbf{V}}_\mathbf{i} = (\mathbf{x}_i)_{\Psi_i\in\hat{\mathcal{Y}}}$. We assume that we model the shape variation locally by a multi variate Gaussian with axes along the principal components of the distribution. Furthermore, without loss of generality we can consider that we can derive a covariance matrix Σ for the position variation of each subset of landmarks in the shape after Procrustes alignment. The alignment discards the influence of global displacements of the local landmark configuration.

For a landmark $\mathbf{x}_i$ not in the sparse representation the reconstruction can be formulated in the following way: We choose the l nearest neighbors of Ψ_i in $\hat{\mathcal{Y}}$ (i.e. the landmarks, that exhibit the highest coherence of behavior in the training set and are part of the sparse representation). The use of the closest neighbors in the shape map ensures a reconstruction based on the functionally closest related landmarks as opposed to spatial neighbors. The model learnt from the training data for this sub-set $\mathbf{V_{i,j}}'$ of landmarks or *reconstruction kernel* comprises a mean μ and a covariance matrix Σ. The shape vector is partitioned into the observed part $\mathbf{V_{i,j}}^a$ of the sparse representation and the missing part $\mathbf{V_{i,j}}^m$ (i.e. one or several missing landmarks). Accordingly we partition the covariance matrix into sub matrix corresponding to the observed values or coordinates Σ^{aa}, and the submatrix corresponding to the missing values Σ^{mm}, the submatrix describing their relation is $\Sigma^{am} = \Sigma^{ma\top}$,

$$\Sigma = \begin{pmatrix} \Sigma^{aa} & \Sigma^{am} \\ \Sigma^{am\top} & \Sigma^{mm} \end{pmatrix}. \tag{9}$$

Now we can estimate the values of the remaining shape landmarks by a linear regression model:

$$\mathbf{V_{i,j}}^m = \mu^m + (\mathbf{V_{i,j}}^a - \mu^a)\mathbf{B} + e, \mathbf{B} = \Sigma^{aa-1}\Sigma^{am}. \quad (10)$$

$\mathbf{B}$ is the regression matrix, $\mathbf{X}^m$ is the conditional maximum likelihood estimate of the missing part of the shape vector, and e is a residual error. See [18] for a concise explanation of imputation. Therefore in the case of a linear model, the local sparse reconstruction function is given by equation (11). It allows to reconstruct the entire object shape from the sparse representation, while using the relations learnt from the training set to define local reconstruction kernels in the map $\mathcal{S}$.

$$\mathbf{V'_{i,j}} \mapsto \begin{pmatrix} \mathbf{V'_{i,j}} \\ \\ \mathbf{V_{i,j}}^m \end{pmatrix} \quad (11)$$

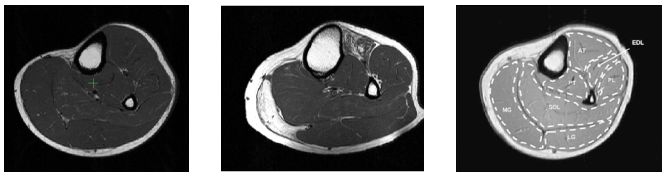
2.5. Inference from new data

Let us now consider a new data set where the goal is to determine the position of the object being modeled. Such an inference process often involves the definition of an objective function that seeks for an admissible solution being supported from the observations. In a standard shape model inference approach, the positions of landmarks in new data are estimated by an energy minimization that involves both shape prior and appearance costs.
The search with the sparse model representation $\hat{\mathbf{V}}$, the according reconstruction function P, and the appearance models $(b_i)_i = 1, \ldots, N$ for each landmark is performed in an iterative manner. Based on a coarse initialization the landmark positions of $\hat{\mathbf{V}}$ are updated according to the appearance model. For each landmark the position with highest probability with regard to a local texture patch is chosen. Then the shape is constraint by either a local or global statistical shape constraint. In our work we use a multivariate Gaussian. However, alternatives, like spherical wavelets [16], or elasticity based constraints [19] can be utilized in a similar manner. After convergence the entire shape $\mathbf{V}$ is reconstructed from $\hat{\mathbf{V}}$ by the sparse reconstruction function.

3. Experimental Validation

3.1. Application Context

In medical image analysis often the assumption of consistent global region statistics is violated within organs. The work was motivated by the highly heterogeneous distribution of visual information in muscle MRI data (Fig. 4). Standard segmentation methods fail, due to the ambiguous texture and the sparse distribution of salient image information for example within the leg. The majority of work on anatomy segmentation is focused on brain, liver or heart

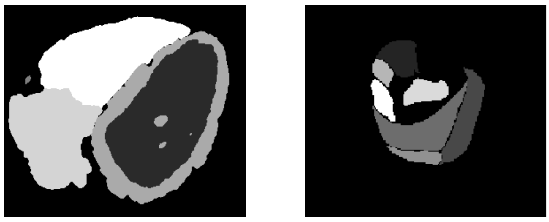


(a) (b) (c)

Figure 4. MRI data of calf muscles: (a) healthy (b) and unhealthy case, (c) manual expert annotation of individual muscles.

data [14, 1]. They are often based either on the detection of prominent edges between organs or on the separation of intensities due to the reflectance properties of the different tissues in the human body. The segmentation of individual muscles within a muscle compound poses new challenges to automatic segmentation systems. The sparse distribution of regions where image information allows for a reliable separation of neighboring substructures, makes the use of prior shape knowledge mandatory, and motivates the development of models, that make optimal use of statistical models of shape and appearance acquired during training.

In Fig 4, magnetic resonance imaging (MRI) slices of the human calf are depicted [10]. The distribution of reliable image information at the boundaries between individual muscles is un-even, and parts can only be estimated from prior shape knowledge. Since this distribution is fairly consistent over a population it can be learned, and integrated in a sparse model that makes optimal use of both the shape prior and the image information.

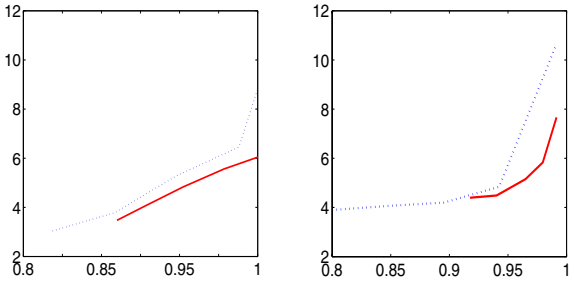


(a) (b)

Figure 5. Standard reference segmentation of respectively (a) papillary muscles of the left ventricle and (b) human calf muscle.

In [8] a hybrid image-based modeling approach is presented, with a predefined reference model and using a free-form deformation approach to morph a high resolution model to fit a low-resolution model created from MR data. The Statistical Shape Model (SSM) proposed by Cootes [5] were used to capture and represent the variation in shape of a set of training sets. From a set of training data the mean shape and its most significant modes of variation are determined. In [9] two-dimensional slices where utilized, to represent the shapes, and a sub-set of these slices was chosen by means of a criterion capturing image and model support.

The sparse representation of shapes presented is based on a finite set of landmarks, that can be identified repeatedly on different examples of an anatomical structure. The use of landmarks extends this concept to the differentiation between arbitrary regions of on the object surface.



(a) Muscle Data (b) Heart Data

Figure 6. Reconstruction: accuracy of the shape reconstruction with different sparsity levels: uniform sub-sampling vs. weighted sub-sampling on muscle and heart data.

3.2. Experimental Set-up

To evaluate the performance of the proposed method we report experiments on two data sets:

1. A set of 4 calf muscles of healthy control patients: for each volume 20 slices of 7mm of thickness, and with voxel spacing of 0.5859x0.5859x7 mm were acquired with a 1.5 T Siemens scanner. Standard of reference annotation by clinical experts for the Medial Gastrocnemius (MG) muscle, was available (Fig. 5). Correspondences for a set of 895 landmarks on the muscle surfaces were obtained by an MDL optimization.

2. A set of 24 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 726 control points for the left ventricle was available.

For both data sets we evaluated the reconstruction and search behavior of sparse models. To assess the shape representation of the sparse sampling, we sub-sampled the shapes with landmarks either evenly distributed in the real space, or evenly distributed in the shape map, while neglecting appearance. The goal is to understand how the sparse sampling based on the density in the shape map affects the reconstruction of missing landmarks.

To evaluate the search behavior we compared sparse shape models with a standard shape model search in an active shape model manner, based on an even sampling of the object surface, and gradients in the volumes.

	M1	M2	M3	M4
Standard model:	34.77	31.56	49.06	35.74
Sparse model:	4.61	13.41	7.65	12.46

Table 1. Calf muscle segmentation: landmark error after finished search standard model, and a sparse model for 4 example data.

3.3. Results

Reconstruction results are shown in Fig. 6. With an equal ratio of missing landmarks (X-axis) the sparse sampling based on the shape map consistently outperforms uniform sub-sampling in the object space. The advantage becomes more pronounced with very high ratios of missing landmarks that hae to be reconstructed. This indicates that a high amount of relevant information can be captured in a small sub-set of landmarks, when the modeling relations between them are considered by means of the shape map.

The search behavior of sparse models was evaluated on both data sets. A sparse representation was built based on both shape model and image support. Models were initialized with minimal overlap to the target shape, and the accuracy of the final result was quantified by means of the mean landmark error between standard of reference annotation and search result. The sparse model was able to recover the shape with superior accuracy. In Tab. 1 mean landmark errors after search convergence for standard shape models, and sparse shape models are reported. In the muscle data the standard search approach failed due to the ambiguous texture in large regions of the target shape. In Fig. 7 examples for standard and sparse model search are depicted. An interesting observation was that in the case of calf muscle image support and diffusion map density gave complementary distributions. That indicates, that it is worthwhile to use both informations for the representation building but raises the question of an appropriate weighting, and its dependence on the data variability. This will be subject of ongoing research on more exhaustive data sets. For the heart data, the search was initialized with minimum overlap. Standard search results in a mean error of 20.86 voxels, while sparse models obtain a mean landmark error of 6.43 voxels.

4. Conclusion

In this paper we propose a knowledge-based segmentation framework. In contrast to existing approaches the method uses the statistical shape modeling and texture behavior in a training set to derive a sparse representation and reconstruction mechanism. It adapts to heterogeneous distributions of redundancy in the shape variation, and sparsely distributed distinctive texture in the data. During search for structures in new data, the model is used with a patch based local appearance representation to locate and segment objects. We report experimental results on two complex data

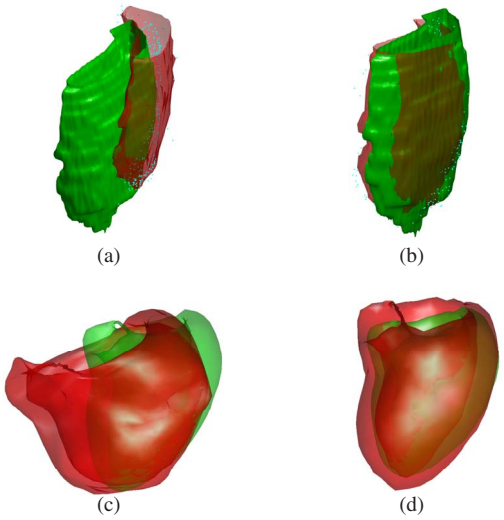


Figure 7. Model search result for MRI calf (upper row) and heart muscle (lower row) data, green: standard of reference segmentation, red: search results for a. and c. standard gradient search approach, and uniform sampling, b. and d. sparse shape models.

sets, and compare the method with existing approaches. The proposed representation can by employed with other shape modeling and search methods. The formulation based on model compactness makes a transfer of the shape behavior mapping to other models straightforward. Future work will focus on the determination of the sparse model complexity, that takes the limited training data into account to estimate a feasible number of parameters, and an integration with model learning approaches, that learn the locations in non-annotated data in a weakly- or unsupervised manner.

5. Acknowledgments

The authors are highly grateful to Dr. Jean François Deux and Pr. Alain Rahmouni from Henry Mondor Universitary Hospital for providing data and ground truth segmentation, and to Ben Glocker for providing us with deformable registration application .

References

[1] Y. Bizais, C. Barillot, and R. D. Paola. *Information Processing in Medical Imaging*. Computational Imaging and Vision. Kluwer Academic Publishers, 1995. 5

[2] F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE TPAMI*, 11(6):567–585, 1989. 1

[3] F. R. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997. 3

[4] R. R. Coifman and S. Lafon. Diffusion maps. *Appl. Comput. Harmon. Anal.*, 21:5–30, 2006. 3

[5] T. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE TPAMI*, 23(6):681–685, 2001. 1, 5

[6] 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. 1

[7] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor. 3d statistical shape models using direct optimisation of description length. *ECCV*, 3:3–20, 2002. 2, 3

[8] J. Fernandez, P. Mithraratne, S. Thrupp, M. Tawhai, and P. Hunter. Anatomically based geometric modelling of the musculo-skeletal system and other organs. *Biomech Model Mechanobiol*, 2:139–155, 2004. 5

[9] C. Florin, N. Paragios, G. Funka-Lea, and J. Williams. Liver segmentation using sparse 3d prior models with optimal data support. *IPMI'07*, pages 38–49, 2007. 5

[10] C. Galban, S. Maderwald, K. Uffmann, and M. E.Ladd. A diffusion tensor imaging analysis of gender differences in water diffusivity within human skeletal muscle. *NMR in Biomedicine*, 2005. 5

[11] H. Lamecker, M. Seebass, H. Hege, and P. Deuffhard. A 3d statistical shape model of the pelvic bone for segmentation. In *Proc. of SPIE 04*, pages 1341–1351, 2004. 1

[12] 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. 2

[13] G. Langs and N. Paragios. Modeling the structure of multivariate manifolds: Shape maps. In *Proc. of CVPR'08*, 2008. 3

[14] T. McInerney and D. Terzopoulos. A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4d image analysis. In *Computerized Medical Imaging and Graphics*, volume 19(1), pages 69–83, 1995. 5

[15] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: A survey. In *Medical Image Analysis 1*, pages 91–108, 1996. 1

[16] D. Nain, S. Haker, A. Bobick, and A. Tannenbaum. Shape-driven 3d segmentation using spherical wavelets. In *Proc. of MICCAI'06*, 2006. 1, 5

[17] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978. 3

[18] T. Schneider. Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate*, 14:853 – 871, 2001. 5

[19] M. Taron, N. Paragios, and M. P. Jolly. From uncertainties to statistical model building and segmentation of the left ventricle. In *Proc. of MMBIA'07*, 2007. 2, 5

[20] T.Deschamps and L.Cohen. Fast extraction of minimal paths in 3d images and application to virtual endoscopy. In *Medical Image Analysis*, volume 5, pages 281–299, December 2001. 1

[21] Z. Tu and S. Zhu. Image segmentation by data-driven markov chain monte carlo. *TPAMI*, 24:657–673, 2002. 1

[22] C. S. Wallace and D. L. Dowe. Minimum message length and Kolmogorov complexity. *The Computer Journal*, 42(4):270–283, 1999. 3