Segmenting Cardiopulmonary Images Using Manifold Learning with Level Sets

Qilong Zhang and Robert Pless

Department of Computer Science and Engineering,
Washington University, St. Louis, MO, 63130, USA
{zql, pless}@cse.wustl.edu

Abstract. Cardiopulmonary imaging is a key tool in modern diagnostic and interventional medicine. Automated analysis of MRI or ultrasound video is complicated by limitations on the image quality and complicated deformations of the chest cavity created by patient breathing and heart beating. When these are the primary causes of image variation, the video sequence samples a two-dimensional, nonlinear manifold of images. Non-parametric representations of this image manifold can be created using recently developed manifold learning algorithms. For automated analysis tasks that require segmenting many images, this manifold structure provides strong new cues on the shape and deformation of particular regions of interest. This paper develops the theory and algorithms to incorporate these manifold constraints within a level set based segmentation algorithm. We apply our algorithm, based on manifold constraints to the problem of segmenting the left ventricle, and show the improvement that arises from using the manifold constraints.

1 Introduction

Images of the chest cavity of a particular patient vary due to the imaging geometry, the permeation of contrast agents through different tissues, noise and deformation caused by the patient's breathing and heartbeat. Accounting for the deformation is vital to many image analysis tasks – this is underscored by the many diagnostic protocols that use gated MR-imaging to minimize motion of the heart, and/or held-breath protocols which ask patients to minimize breathing motions.

When many images are taken with the same imaging geometry of the same patient, these images are samples of a manifold with two degrees of freedom in principle: the phase of the heartbeat and the phase of the breathing cycle. When sufficiently many images are available, manifold learning algorithms, (typified by Isomap [1], Locally Linear Embedding (LLE) [2], and Semidefinite Embedding (SDE) [3]), create nonparametric representations of low-dimensional nonlinear manifolds.

The contribution of this paper is to develop an approach that first learns the manifold structure of the images of a particular patient, then exploits this structure to improve segmentation. For segmentation of the left ventricle, for example,

Y. Liu, T. Jiang, and C. Zhang (Eds.): CVBIA 2005, LNCS 3765, pp. 479–488, 2005.

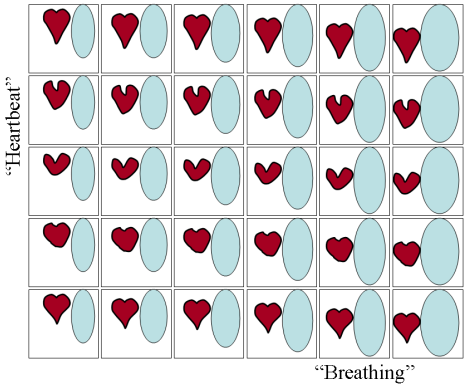

Fig. 1. A cartoon of the manifold structure of a cardiopulmonary video. Manifold learning techniques can automatically parameterize a video sequence by the position of each image on this cardiopulmonary manifold. This manifold structure provides additional cues for segmenting multiple images — for example, motion along the breathing axis simply translates the heart image, while motion along the heartbeat axis deforms the heart shape (with minimal global translational motion). These constraints are more specific and therefore stronger than temporal smoothness constraints. This paper develops methods to enforce these manifold based constraints.

the manifold structure gives strong cues about the shape changes between manifold neighbors. Figure 1 depicts the constraints in the cardiopulmonary manifold. In particular, variations in the breathing phase lead to an approximately uniform translation of the heart, corresponding to a rigid translation of the (2D) shape segments. Alternatively, changes in the heartbeat phase lead to variations in the shape of the heart, but, largely, not its position. These more specific constraints provide stronger cues than those available from just the temporal order of the original video, in which consecutive images often exhibit both a translation and a non-rigid deformation.

For segmentation of MR-imagery, several representation tools allow the description of shapes within an image, and support algorithms that automatically fit these shapes to image data. One such tool is level sets [4,5], which represents 2D shapes as the zero-crossing of a surface. The standard evolution equations which drive the adaptation of the surface to the image data have very natural modifications which allow the level set to enforce manifold based-constraints. In preliminary experimental results, we find that these additional constraints that the manifold imposes on the level set evolution allows segmentation of the left ventricle in images that are too low contrast to support single image segmentation.

The following section gives a very brief background in manifold learning and highlights previous work specializing these algorithms for biomedical applications. This is followed by a description of the standard Level Set framework for object segmentation. Section 3 describes a derivation of new level set evolution

equations that enforce the manifold-based constraints on shape changes between images. We conclude with experiments on a low-contrast cine-MRI sequences.

2 Background and Previous Work

This work integrates ideas from level set segmentation and manifold learning. To our knowledge, it is novel to combine these approaches. In order to ground our later presentation, we first introduce, very briefly, some recent research in the use of level sets in biomedical image analysis and an overview of manifold learning.

2.1 Manifold Learning

Image data can be naturally represented as points in a high dimensional data space (one dimension for each pixel). When the set of images has lower *intrinsic* dimensionality, the data set can be mapped onto a lower dimensional space. Principle Component Analysis (PCA) [6] and Independent Component Analysis (ICA) [7] are two algorithms that represent data as linear combinations of basis vectors — the coefficients that specify each image define a linear low-dimensional embedding of the data set.

However, often the number of basis images required to a linearly reconstruct a data set is much larger than the degrees of freedom in the process that generates the images. This has led to a number of methods seeking to parameterize low-dimensional, nonlinear manifolds. These methods measure local distances or approximate geodesic distances between points in the original data set, and seek low-dimensional embeddings that preserve these properties. Isomap [1] extends classic multidimensional scaling (MDS) by substituting an estimate of the geodesic distance along the image manifold for the inter-image Euclidean distance. LLE [2] attempts to represent the image manifold locally by reconstructing each image as weighted combination of its neighbors. SDE [3] applies semidefinite programming to learn kernel matrices which can be used to create isometric embeddings. These algorithms, and others [8,9,10] have been used in various applications, including classification, recognition, tracking, and to a limited extent, biomedical image analysis [11].

2.2 Isomap Embedding of Cardiopulmonary Manifolds

The Isomap procedure for dimensionality reduction starts by computing the distance between all pairs of images (using some distance function such as SSD pixel intensities). Then, a graph is defined with each image as a node and undirected edges connecting each image to its k-closest neighbors (usually choosing k between 5 and 10). A complete pair-wise distance matrix is calculated by solving for the all-pairs shortest paths in this sparse graph. Finally, this complete distance matrix is embedded into some low dimension by solving an Eigenvalue problem (Multidimensional Scaling (MDS) [12]). The dimensionality embedding can be chosen as desired, but ideally is the number of degrees of freedom in

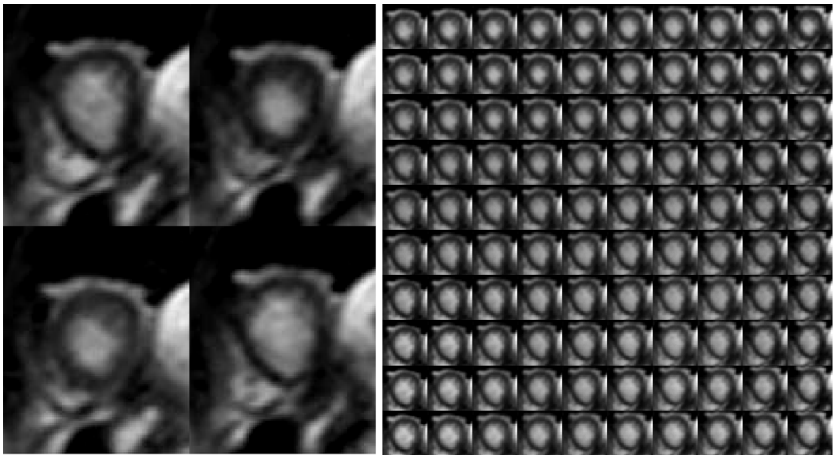

Fig. 2. (left) four images from a 200 image cardiopulmonary cine-MRI sequence of the heart. Note the variation both in the shape of the left ventricle (the white blob roughly centered in the image) and the position of the heart (shifting vertically). (right) A 2D manifold was defined with Isomap, using a Gabor filter-based image distance function. The 2D embedding is interpolated to give a regular sampling of the image manifold, and Section 3 modifies the level set approach to segment images over the entire manifold simultaneously.

the image set, in our case 2 (the two intrinsic dimensions of variability are the heartbeat and breathing).

Previous work that applies manifold learning to biomedical image analysis suggests modifying Isomap to use image distance functions other than pixel intensity differences [13,14]. For data sets with deformable motion, the suggested distance function is computed as the phase difference of local complex Gabor filters:

$$||I_1 - I_2||_{motion} = \sum_{x,y} \Psi(G_{(\omega,V,\sigma)} \otimes I_1, G_{(\omega,V,\sigma)} \otimes I_2) + \Psi(G_{(\omega,H,\sigma)} \otimes I_1, G_{(\omega,H,\sigma)} \otimes I_2)$$

where $G_{(\omega,\{V|H\},\sigma)}$ is defined to be the 2D complex Gabor filter with frequency ω, oriented either vertically or horizontally, with σ as the variance of the modulating Gaussian, and Ψ returns the phase difference of the pair of complex Gabor responses above some threshold τ; we choose τ to be the 50-th percentile filter magnitude. Figure 2 gives 4 example images of the heart, illustrating both the non-rigid and the rigid deformations. The Isomap embedding computes a 2D coordinate for each original image.

An even sampling of this manifold simplifies the numerical implementation of the level set segmentation in the subsequent sections. However, the given image sequence may not be evenly distributed in the manifold space. Ideally, it is desirable to have a continuous image function f to describe all possible cardiopulmonary images. One may interpolate the image function f locally by

fitting thin-plate smoothing spline to the given images $\{\mathrm{I}_i \mid i = 1 \cdots n\}$ and their associated manifold position $\{(u_i, v_i) \mid i = 1 \cdots n\}$, such that $f(x, y, u, v)$ minimizes the following weighted sum:

$$(1-p)\sum_{i=1}^{n} |I_i(x,y) - f(x,y,u_i,v_i)|^2 \ldots$$
$$+ p \int \left|\frac{\partial^2 f(x,y)}{\partial u^2}\right|^2 + 2\left|\frac{\partial^2 f(x,y)}{\partial u \partial v}\right|^2 + \left|\frac{\partial^2 f(x,y)}{\partial v^2}\right|^2 \, dudv, \tag{1}$$

where p is a the smoothing parameter. We use the Matlab toolbox function *tpaps*, which chooses this smoothing parameter "in an ad hoc fashion in dependence on the [data]" [15].

In the next section, we consider how to exploit this manifold structure, in order to assist a level set function to track the left ventricle of a beating heart.

3 Level Set Segmentation

This section refines a standard framework for level set segmentation. The presentation includes background material through Equation 5, after which we extend the approach to segment all images in the data set simultaneously, and enforce additional constraints from the manifold structure.

In n-dimension space Ω, we define the evolving hypersurface C as the boundary $\partial\Omega$ of regions of interest. We call Ω^- the inside of C and Ω^+ the outside of C. For the cases of image segmentation, on approach is to define the contour as an energy minimization problem (the following presentations follows [16]):

$$\begin{aligned} E(c_1, c_2, C) = \mu \cdot Length(C) + \nu \cdot Area(\Omega^-) \\ + \lambda_1 \int_{\Omega^-} |f(x,y) - c_1|^2 dx \; dy \\ + \lambda_2 \int_{\Omega^+} |f(x,y) - c_2|^2 dx \; dy, \end{aligned} \tag{2}$$

where c_1 and c_2 are constants depending on C and are usually the average of image intensity $f(x, y)$ in the region Ω^- and the outside Ω^+, respectively. All parameter settings, such as λ_1, used in our experiments are listed in Section 4.

In problems of curve evolution, the level set method has been used extensively. Using the level set formulation, the boundary C is represented by the zero level set of a Lipschitz function $\phi : \Omega \to \mathbb{R}$, such that:

$$\begin{cases} C = \partial\Omega = \{(x,y) \in \Omega : \phi(x,y) = 0\} \\ \Omega^- = \{(x,y) \in \Omega : \phi(x,y) < 0\} \\ \Omega^+ = \{(x,y) \in \Omega : \phi(x,y) > 0\} \end{cases}$$

Using the Heaviside and Dirac delta functions with a smoothed approximation of finite width ϵ:

$$H_\epsilon(\phi) = \frac{1}{2}\left[1 + \frac{2}{\pi}\arctan\left(\frac{\phi}{\epsilon}\right)\right], \quad \delta_\epsilon(\phi) = \frac{d}{d\phi} H_\epsilon(\phi) \tag{3}$$

the energy functional (2) can be written to evaluate the level set function ϕ on the domain Ω:

$$\begin{aligned} E(c_1, c_2, \phi) = \mu \int_\Omega \delta_\varepsilon(\phi(x,y))|\nabla\phi(x,y)|dx\ dy + \nu \int_\Omega H_\varepsilon(\phi(x,y))dx\ dy \\ + \lambda_1 \int_\Omega |f(x,y) - c_1|^2(1 - H_\varepsilon(\phi(x,y)))dx\ dy \\ + \lambda_1 \int_\Omega |f(x,y) - c_2|^2 H_\varepsilon(\phi(x,y))dx\ dy \end{aligned} \tag{4}$$

Using the calculus of variations, one can recover the following evolution equation which incorporates an artificial time parameter t and converges to minimize $E(c_1, c_2, \phi)$:

$$\frac{\partial\phi}{\partial t} = \delta_\varepsilon(\phi)\left[\mu\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \nu + \lambda_1(f - c_1)^2 - \lambda_2(f - c_2)^2\right] \tag{5}$$

To complete the definition of this evolution, we need to additional define the starting condition, $\phi_0(x, y)$. Section 4 describes the starting conditions used in our experiments. The next section illustrates how to incorporate manifold constraints on the level set solution, including the additions to the energy function E and the solution for the corresponding evolution equation.

3.1 Level Set Segmentation on Image Manifolds

For cardiopulmonary image sequences, the images vary in principle depending on their cardiac phase u and pulmonary phase v — the two degrees of freedom that parameterize the manifold shown in Figure 2. As described in Section 2.2, we use Isomap to automatically parameterize all images, and interpolate the result to generate evenly spaced samples of the image manifold $f(x, y, u, v)$. Accordingly, the contour C that we seek is also a function of u and v, and our goal is to describe C implicitly by the level set function ϕ in 4-dimension space Ω. Thus, a given cardiopulmonary image sequence specifies this contour by extending the energy functional (2) to 4-dimension space:

$$\inf_{c_1,c_2,\phi} E_1(c_1, c_2, \phi), \text{ where } \phi: \mathbb{R}^4 \to \mathbb{R} \tag{6}$$

But the manifold dimensions also correspond to specific kinds of deformation. The breathing of the patient results, approximately, in a translation of the heart. Therefore, we expect the variation of ϕ in the v direction to be a uniform translation. That is, the energy functional change $\frac{\partial\phi}{dv}$ should be consistent with a uniform translation. This induces a level set corollary to the classic optic flow constraint equation [17]:

$$\frac{\partial\phi}{\partial x}\omega_x + \frac{\partial\phi}{\partial y}\omega_y + \frac{\partial\phi}{\partial v} = 0 \tag{7}$$

where $(\omega_x,\ \omega_y)^\top$ is the velocity vector that is constant over any given image, but may vary for different values of u and v.

On the other hand, varying images along the other axis of the image manifold, deformations due to the cardiac cycle lead to image variation with minimal overall translation. For the special case of deformation caused by (non-uniform) heart expansion and contraction, we can express the constraint as:

$$\frac{\partial \phi}{\partial u} = \omega_u \tag{8}$$

where ω_u is constant over the region of the heart for any given u and v. This constraint enforces the condition that moving along the "heartbeat" axis simply adds or subtracts a constant value of the Level Set function ϕ, and therefore enforces that the shape either expands or shrinks.

Enforcing these constraints is natural within the level set framework; both lead naturally to new terms in the evolution of the ϕ function. Computationally, at each time step of the temporal evolution ϕ, we compute three parameters for each sample manifold image (u, v). First, we compute the least-squares estimates of the vector $(\omega_x,\ \omega_y)^\top$ that corresponds to global motion as we move from one image to another in the u direction on the image manifold. Second, we compute the constant ω_u that defines the change to ϕ the best corresponds to the expansion or contraction of the shape as we move from one image to another on the manifold along the v direction.

That is, for a particular u, v value, we compute first $\omega_x(u, v), \omega_y(u, v)$ by calculating $\frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}$, and $\frac{\partial \phi}{\partial u}$ over the entire image, and solving the resulting linear system (from Equation (7)). Second, we compute $\omega_u(u, v)$ at a particular point u, v as the mean value (over all x, y) of $\frac{\partial \phi}{\partial u}$.

Once we have computed $\omega_x(u, v), \omega_y(u, v)$, and $\omega_u(u, v)$, we can write the motion constraints as an energy functional:

$$\begin{aligned} E_2(\phi) = \eta_1 \int_\Omega \left(\frac{\partial \phi}{\partial x}\omega_x + \frac{\partial \phi}{\partial y}\omega_y + \frac{\partial \phi}{\partial v} \right)^2 d\Omega \\ + \eta_2 \int_\Omega \left(\frac{\partial \phi}{\partial u} - \omega_u \right)^2 d\Omega \end{aligned} \tag{9}$$

where η_1 and η_2 are blending parameters. The first term enforces rigid changes in shape by penalizing regions of ϕ where the x, y, v derivatives are not consistent with the translation motion, and the second term penalizes the overall mean translational motion of the heart, which is minimal when motion is caused only by the heartbeat.

Solving for the evolution equation such that ϕ minimizes $E_2(\phi)$ gives:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = 2\eta_1 \left(\frac{\partial^2 \phi}{\partial x^2}\omega_x^2 + \frac{\partial^2 \phi}{\partial y^2}\omega_y^2 + \frac{\partial^2 \phi}{\partial v^2} \right. \\ \left. + 2\frac{\partial^2 \phi}{\partial x \partial y}\omega_x\omega_y + 2\frac{\partial^2 \phi}{\partial x \partial v}\omega_x + 2\frac{\partial^2 \phi}{\partial y \partial v}\omega_y \right) + 2\eta_2 \left(\frac{\partial^2 \phi}{\partial u^2} - \frac{\partial \omega_u}{\partial u} \right). \end{aligned} \tag{10}$$

One can integrate this motion constraint module (9) into the previously defined energy term (6):

$$E(c_1, c_2, \Psi, \phi) = E_1(c_1, c_2, \phi) + E_2(\Psi, \phi). \tag{11}$$

Given an initial level set function ϕ_0, we minimize the above functional (11) by iterating two steps, first using the current estimate of ϕ to estimate c_1, c_2 and solving for $\omega_x(u,v), \omega_y(u,v)$, and $\omega_u(u,v)$, and then evolving ϕ by Equations (5) and (10).

4 Example Application

In cardiovascular imagery, an important application is to measure the dynamic behavior of the human heart, especially the left ventricle [18]. Here we define all

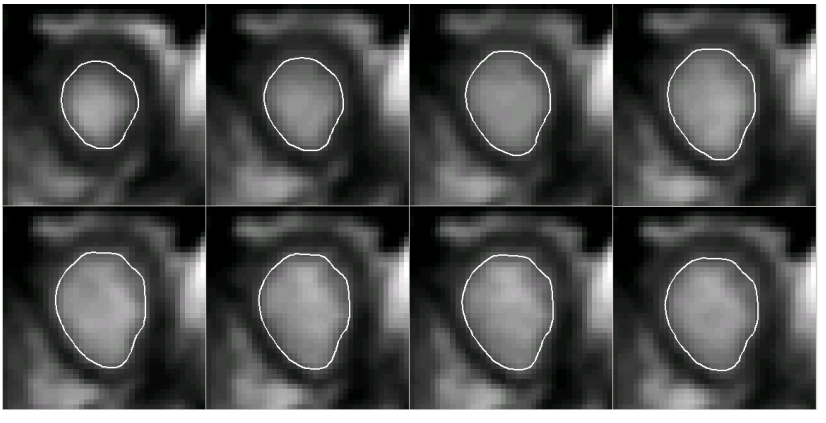

Fig. 3. Segmentation examples of cine-MRI images (from the same data set as shown in Figure 2)

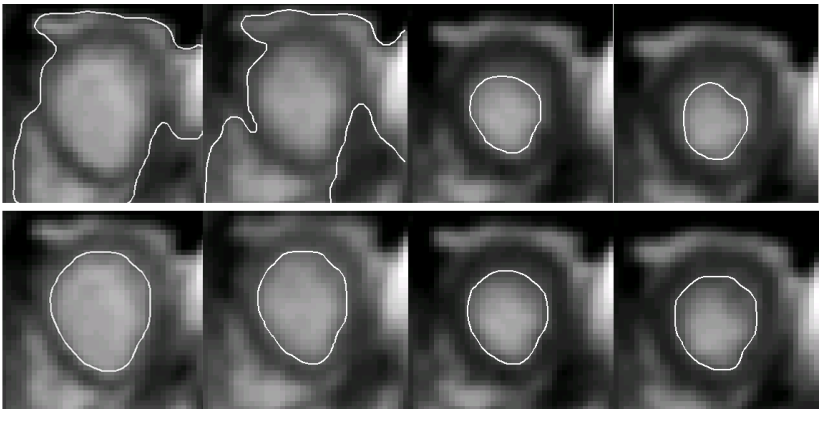

Fig. 4. A comparison of single image based segmentation (top) and the simultaneous solution for all image using the manifold constraints. In the left two images, the single image solution fails because of low contrast, on the right manifold based solution differs and draws a perceptually more reasonable boundary.

the parameters of the algorithm described in the last section and show preliminary results for segmenting a 200 frame cine-MRI sequence (the same data set as shown in Figure 2), using the manifold based constraints.

Our version of the level set algorithm has several parameters. The standard level set parameters λ_1 and λ_2 determine the importance of matching the intensity estimate for the inside and outside of the contour, and are set to $\lambda_1 = \lambda_2 = 1$. The parameter μ penalizes long contour curve lengths, and is set to $\mu = 0.075$. We set the area penalty parameter $\nu = -0.1$, which tends to make the shape grow. We use $\epsilon = 1$ for approximating the Heaviside function H_ϵ and Dirac function δ_ϵ (in Equation (3)).

There are also parameters specific to our modification of the level set function. Our experimental data set has 200 images, and we constructed the image function $f(x, y, u, v)$ of size $40 \times 40 \times 10 \times 10$ to regularly sample the manifold. The blending parameters η_1 and η_2 determine the importance of manifold constraints, and are set to $\eta_1 = \eta_2 = 0.1$. The initial level set $\phi_0(x, y, u, v)$, for each image, is defined by the signed distance function to the circle of radius 6 centered at the image center, with points inside the circle having a negative value. Figure 3 show examples of the segmentation solution for eight consecutive frames using these parameters.

The manifold constraints are most important for images that are especially low contrast or noisy. Figure 4, gives examples of images where the manifold based solution differs significantly from the single image solution. In the first two cases, the manifold constraints show a significant improvement — the single image solutions are wrong because of insufficient contrast. The last two cases show segmented shape boundaries that are different than the single image segmentation, and which may reflect more accurately the correct boundary, although it is difficult to quantify the improvement.

5 Summary and Discussion

This work presents preliminary efforts towards incorporating manifold learning as a tool to provide additional constraints for segmenting cardioplumonary images. This approach can be applied to any application domain for which there is a known manifold structure to the data, and may be extended also to other computational shape representation tools (such as snakes). Furthermore, we believe that many algorithms may be improved through better understanding and exploitation of non-linear image manifold learning algorithms, and tight integration of these with classical analysis tools.

References

1. Tenenbaum, J.B., de Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. Science **290** (2000) 2319–2323
2. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. Science **290** (2000) 2323–2326

3. Weinberger, K.Q., Saul, L.K.: Unsupervised learning of image manifolds by semidefinite programming. In: Computer Vision and Pattern Recognition. (2004)
4. D.Mumford, Shah, J.: Optimal approximation by piecewise smooth function and associated variational problems. In: Commun. Pure Appl. Math. Volume 4. (1989) 577–685
5. Osher, S., Fedkiw, R.: The Level Set Method and Dynamic Implicit Surfaces. Springer-Verlag, New York (2003)
6. Jolliffe, I.T.: Principal Component Analysis. Springer-Verlag (1986)
7. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. John Wiley and Sons (2001)
8. Donoho, D.L., Grimes, C.: Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data. PNAS **100** (2003) 5591–5596
9. Brand, M.: Charting a manifold. In S. Becker, S.T., Obermayer, K., eds.: Advances in Neural Information Processing Systems 15. MIT Press, Cambridge, MA (2003) 961–968
10. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Advances in Neural Information Processing Systems. (2002)
11. Lim, I.S., Ciechomski, P.d.H., Sarni, S., Thalmann, D.: Planar arrangement of high-dimensional biomedical data sets by isomap coordinates. In: Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems (CBMS 2003). (2003) 50–55
12. Borg, I., Groenen, P.: Modern Multidimensional Scaling: Theory and Applications. Springer-Verlag (1997)
13. Pless, R.: Differential structure in non-linear image embedding functions. In: Articulated and Nonrigid Motion. (2004)
14. Souvenir, R., Pless, R.: Isomap and non-parametric models of image deformation. In: In Proc. IEEE Workshop on Motion and Video Computing, Breckenridge, CO (2005)
15. Matlab version 6.5.0 tpaps documentation. Mathworks Inc. (2002)
16. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. on Image Processing **10** (2001) 266–277
17. Horn, B.K.P.: Robot Vision. McGraw Hill, New York (1986)
18. Paragios, N.: A variational approach for the segmentation of the left ventricle in cardiac image analysis. International Journal of Computer Vision, (2002) 345–362