

# Nonlinear Embedding towards Articulated Spine Shape Inference Using Higher-Order MRFs<sup>\*</sup>

Samuel Kadoury<sup>1</sup> and Nikos Paragios<sup>2</sup>

<sup>1</sup> Philips Research North America, Briarcliff Manor, NY, USA  
samuel.kadoury@philips.com

<sup>2</sup> Ecole Centrale de Paris, Laboratoire MAS  
GALEN Group, INRIA Saclay, Ile-de-France, France  
nikos.paragios@ecp.fr

**Abstract.** In this paper we introduce a novel approach for inferring articulated spine models from images. A low-dimensional manifold embedding is created from a training set of prior mesh models to establish the patterns of global shape variations. Local appearance is captured from neighborhoods in the manifold once the overall representation converges. Inference with respect to the manifold and shape parameters is performed using a Markov Random Field (MRF). Singleton and pairwise potentials measure the support from the data and shape coherence from neighboring models respectively, while higher-order cliques encode geometrical modes of variation for local vertebra shape warping. Optimization of model parameters is achieved using efficient linear programming and duality. The resulting model is geometrically intuitive, captures the statistical distribution of the underlying manifold and respects image support in the spatial domain. Experimental results on spinal column geometry estimation from CT demonstrate the approach's potential.

## 1 Introduction

Statistical models of shape variability have been successful in addressing fundamental vision tasks such as segmentation and registration in medical imaging. For example, Active Shape and Appearance Models have been used in recovering single object geometries obtained from dense collection of data points. Implicit representations is an alternative formulation to address model-based segmentation while more recently numerous methods based on point distribution models (PDM) and embedding on various geometric spaces (spherical [1]) have been proposed. However, in the case of articulated objects such as the spinal column, model-based segmentation of single objects typically leads to fitting errors when pathologies are present. The result is thereby sensitive to model initialization and therefore limited to the capture range. Simultaneous multi-object inference is often beneficial compared to the separate segmentation of individual objects. In [2], an extension of PDMs was considered for modeling relations between shape constellations using conditional probabilities between 2D-contours. Rigid

---

<sup>\*</sup> Funding provided by Stereos+ (Medicen), INRIA and FQRNT grants.

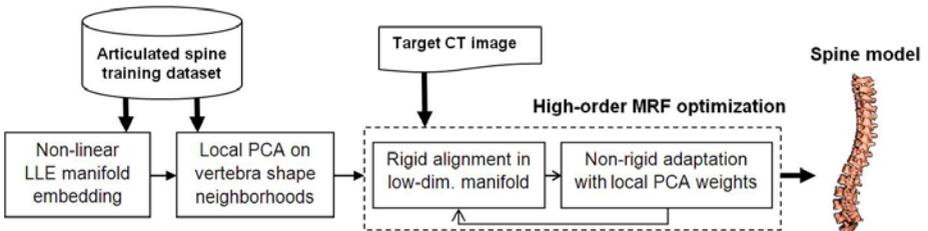
transformations were considered as statistics between part sets to create reconstructed models [3], achieve CT segmentation [4] and multi-modal inference [5].

On the other hand, the high dimensionality and complex non-linear underlying structure unfortunately makes the commonly used linear statistics inapplicable for articulated structures. A manifold learning algorithm of particular interest to this work is locally linear embedding [6]. It maps high-dimensional observation data that are presumed to lie on a nonlinear manifold, onto a single global coordinate system of lower dimensionality. Such a concept reveals the underlying structure of the data which can be used for statistical modeling. Inferring a model from the underlying manifold is a novel concept but far from being trivial. In this work, we tackle the problem of spinal deformity pathologies to model both global statistics of the articulated model and local shape variations of vertebrae based on local measures. We propose a spine inference/segmentation method from CT, where the model representation is optimized through a Markov Random Field (MRF) graph, balancing prior distribution with image data.

We introduce a deformable articulated spine instantiation through a statistical modeling of inter-object transformations. Our method is structured in two parts. The first relates to the creation of a nonlinear manifold embedding of spine articulations which can handle both small and large deformations in a given population. To this end, a novel articulated metric is introduced to create local linear patches. Second, we propose an inference framework using high-order MRF. This graph involves costs related to the imaged data, prior geometrical dependencies and global higher-order cliques. This paper is organized as follows. Section 2 presents the theoretical methodology for the manifold representation of articulated mesh models, while in Section 3 we propose the MRF-based inference framework. In Section 4 we present our evaluation results applied to images of the spinal column and the last section concludes the paper.

## 2 Manifold Embedding of Articulated Spine Models

The method performs an embedding of a training set of annotated vertebra shape constellations into a sub-space which dimensionality corresponds to the domain of admissible variations. Local shape is determined via analysis of variations in patch of the manifold. Fig. 1 illustrates a flowchart of the method.



**Fig. 1.** Flowchart diagram of the proposed spine inference approach.

## 2.1 Representation of Articulated Spine Models

Our spine model  $\mathbf{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_L\}$  consists of an interconnection of  $L$  vertebrae. For each vertebra  $\mathbf{s}_i$ , we recover a triangular mesh with vertices  $\{\mathbf{v}_j^i | j = 1, \dots, V\}$ , where the  $j^{\text{th}}$  vertex corresponds to approximately the same location from one shape to another and  $V$  the number of vertices. Additionally, every  $\mathbf{s}_i$  is annotated with landmarks on each model to rigidly register each object to its upper neighbor. Hence, an articulated deformable model (ADM) is represented by a vector of local inter-vertebral rigid transformations  $A = [T_1, T_2, \dots, T_L]$ . To perform global shape modeling of  $\mathbf{S}$ , we convert  $A$  to an absolute representation  $\mathbf{A}_{\text{abs}} = [T_1, T_1 \circ T_2, \dots, T_1 \circ T_2 \circ \dots \circ T_L]$  using recursive compositions. The transformations are expressed in the local coordinate system (LCS) of the lower vertebra. Center of transformation is the intersection of all 3 vertebral axes, following antero-posterior, cranial-caudal and left-right directions. Rigid transformations described here are the combination of a rotation matrix  $R$ , a translation  $t$  and scaling  $s$ . We formulate the rigid transformation  $T = \{s, R, t\}$  of a triangular mesh model as  $y = sRx + t$  where  $x, y, t \in \mathbb{R}^3$ .

## 2.2 Nonlinear Manifold Embedding of Articulated Spine

Given a training set of  $N$  articulated spine shape models  $\mathbf{S}$  expressed by the absolute vector representation  $\mathbf{A}_{\text{abs}}^i$  of dimensionality  $D$ , we seek their low-dimensional manifold  $\mathcal{M}$  with points  $Y_i, Y_i \in \mathbb{R}^d$  where  $d \ll D$ . The set consisted of pre-reconstructed patient X-rays exhibiting a wide range of deformities (normal to severe). A free-form deformation was applied to obtain vertebral meshes with vertex correspondences. In all, 6 classes of spine deformations were included to represent the underlying population structure. In the sub-cluster corresponding to a pathological population, each individual point of the training set and its neighbours would lie within a locally linear patch on the manifold.

**Nearest neighbor selection.** In our approach, we adopt the intrinsic nature of the Riemannian manifold geometry allowing us to discern between articulated shape deformations in a topologically suited framework. The  $K$  closest neighbours are selected for each point using a distortion metric adapted for geodesic metrics, defined as  $d_{\mathcal{M}}(\mathbf{A}_{\text{abs}}^i, \mathbf{A}_{\text{abs}}^j)$  which estimates the distance of models  $i, j$ . Distance measure can be expressed as a sum of the  $L$  articulation deviations:

$$d_{\mathcal{M}}(\mathbf{A}_{\text{abs}}^i, \mathbf{A}_{\text{abs}}^j) = \sum_{k=1}^L d_{\mathcal{M}}(T_k^i, T_k^j) = \sum_{k=1}^L \|c_k^i - c_k^j\| + d_G(R_k^i, R_k^j) \quad (1)$$

where the canonical representation encodes the intrinsic ( $c$ ) and orientation ( $R$ ) parameters. The first term evaluates intrinsic distances in the  $L_2$  norm evaluating the inter-vertebral translations. The second defines a diffeomorphism  $d_G$  between rotation neighborhoods. We use the concept of geodesics to evaluate the unfolding in the manifold space  $\mathcal{M}$  of orientation vectors in order to estimate how both diverge from the tangent plane.

**Forward and inverse manifold mapping.** The manifold is built by locally linear embedding [6], where the low-dimensional representation of an absolute articulated vector  $\mathbf{A}_{\text{abs}}^i$  is obtained from its neighbors  $\mathbf{A}_{\text{abs}}^j$  for all data points. Hence, a new ADM can be inferred in the embedded  $d$ -space as a low-dimensional data point by finding its optimal manifold coordinates  $Y_i$ . To obtain the articulation vector for a new embedded point in the ambient space, one has to determine the representation in high-dimensional space based on its intrinsic coordinates. The inverse mapping of  $Y_i$  is performed by estimating the relationship between the  $D$ -space and manifold  $\mathcal{M}$  as a joint distribution. The manifold should follow a conditional expectation which captures the overall trend of the data in a local neighborhood of the manifold. Using *Nadaraya-Watson* regression, Gaussian kernels are used to estimate these densities in the conditional expectation setting [7]. By assuming the kernels  $G$  generalizes the expectation such that the observations are defined in terms of a metric  $d_{\mathcal{M}}$  in manifold space:

$$f_{\text{NW}}(Y_i) = \underset{\mathbf{A}_{\text{abs}}^i}{\text{argmin}} \frac{\sum_{j \in \mathcal{N}(i)} G(Y_i, Y_j) d_{\mathcal{M}}(\mathbf{A}_{\text{abs}}^i, \mathbf{A}_{\text{abs}}^j)}{\sum_{j \in \mathcal{N}(i)} G(Y_i, Y_j)} \quad (2)$$

which integrates the distance metric  $d_{\mathcal{M}}(\mathbf{A}_{\text{abs}}^i, \mathbf{A}_{\text{abs}}^j)$  defined in (1) and updates  $f_{\text{NW}}(Y_i)$  using the closest neighbors of point  $Y_i$  in the manifold space. This constrains the regression to be valid for similar data points in its vicinity since locality around  $Y_i$  preserves locality in  $\mathbf{A}_{\text{abs}}^i$ .

### 2.3 Local Vertebra Appearances in the Manifold

The key idea of capturing vertebral shape appearance lies on the assumption that global models, represented in a local neighborhood of  $\mathcal{M}$ , will also manifest similar local geometries. The motivation stems from the fact that global shape deformation belonging to the same class will induce similar local biomechanical patterns, creating morphologically comparable vertebrae. We assume here that vertebra appearances follow a linear distribution within the low-dimensional manifold. Hence, given a data point  $Y_j$  and its  $K$  neighbors, the local shape model  $\mathbf{s}_i$ , representing the  $i^{\text{th}}$  element of the ADM, is obtained by building a particular class of shapes given the set of examples  $\{\mathbf{s}_i^1, \dots, \mathbf{s}_i^K\}$ . We approximate the distribution of the shape using a parameterized linear model by computing the deformation vectors formed for the  $K - 1$  shape samples. We compute the  $n$  eigenvalues and corresponding eigenvectors  $v$  so that a new vertebra  $\mathbf{s}_i^{\text{new}} = \bar{\mathbf{s}}_i + [v_1 \dots v_n][\omega_1 \dots \omega_n]$  can be instantiated where  $\bar{\mathbf{s}}_i$  is the mean shape of the  $K$  neighboring local objects and  $\mathbf{w} = [\omega_1 \dots \omega_n]$  the weight vector. This step warps individual instances to infer new local vertebra models.

## 3 Inference through MRF Optimization

Once an appropriate modeling of spine shape variations is determined, a successful inference between the image and manifold must be accomplished. We describe

here how a new model is deformed. We search the optimal embedded manifold point  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_d)$  of the global spine model. Such a strategy offers an ideal compromise between the prior constraints, as well as the individual shape variation described by the weight vector  $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$  in a localized sub-patch. The energy  $E$  of inferring the model  $\mathbf{S}$  in the image  $\mathcal{I}$  is a function of the set of displacement vectors  $\Delta$  in the manifold space for global shape representation. This involves: (a) a data-related term expressing the image cost and (b) a global prior term measuring deformation between low-dimensional vectors with shape models. The third term represents (c) a higher-order term which is expressed by the reconstruction weights  $\Omega$  for local vertebra modeling. The energy  $E$  can be expressed as the following combination of a global and local optimization:

$$E(\mathbf{S}^0, \mathcal{I}, \Delta, \Omega) = V(\mathbf{Y}^0 + \Delta, \mathcal{I}) + \alpha V(\mathbf{N}, \Delta) + \beta V(\mathbf{H}, \Delta, \Omega). \quad (3)$$

### 3.1 Rigid Alignment of the Spine

The global alignment of the model with the target image primarily drives the deformation of the ADM. The purpose is to estimate the set of articulations describing the global spine model by determining its optimal representation  $\mathbf{Y}^0$  in the embedded space. This is performed by obtaining the global representation using the mapping in (2) so that:  $f_{\text{NW}}(Y_i + \Delta) = f_{\text{NW}}(\{y_1 + \delta_1, \dots, y_d + \delta_d\})$ . This allows to optimize the model in manifold space coordinates while retrieving the articulations in  $\mathcal{I}$ . The global cost can be expressed as:

$$V(\mathbf{Y}^0 + \Delta, \mathcal{I}) = V(f_{\text{NW}}(\{y_1 + \delta_1, \dots, y_d + \delta_d\}), \mathcal{I}). \quad (4)$$

The inverse transform allows to obtain  $\mathbf{A}_{\text{abs}}^i + \mathbf{D}$ , with  $\mathbf{D}$  as deformations in the image space. Since the transformations  $T_i$  are implicitly modeled in the absolute representation  $\mathbf{A}_{\text{abs}}^0$ , we can formally consider the singleton image-related term as a summation of costs associated with each  $L$  vertebra of the ADM:

$$V(\mathbf{A}_{\text{abs}}^0 + \mathbf{D}, \mathcal{I}) = \sum_{i=1}^L V_i(\mathbf{s}_i * (T_i^0 + \mathbf{d}_i), \mathcal{I}) \quad (5)$$

where  $V_i(\mathbf{s}, \mathcal{I}) = \sum_{\mathbf{v}_i \in \mathbf{s}} \mathbf{n}_i^T(\mathbf{v}_i) \nabla \mathcal{I}(\mathbf{v}_i)$  minimizes the distance between mesh vertices of the inferred ADM and gradient image  $\mathcal{I}$  by a rigid transformation. Here,  $\mathbf{n}_i$  is the normal pointing outwards and  $\nabla \mathcal{I}(\mathbf{v}_i)$  the image gradient at  $\mathbf{v}_i$ .

The prior constraint for the rigid alignment are pairwise potentials between neighboring models  $y_i$  such that the difference in manifold coordinates is minimal with regards to a prior distribution of neighboring distances  $P$ :

$$\alpha V(\mathbf{N}, \Delta) = \alpha \sum_{i \in G} \sum_{j \in \mathcal{N}(i)} V_{ij}(y_i^0 + \delta_i, y_j^0 + \delta_j, P). \quad (6)$$

This term represents the smoothness term of the global cost function to ensure that the deformation  $\delta_i$  applied to point coordinates are regular, with  $V_{ij} = (0, 1)$  a distance assigning function based on the distances to  $P$ .

### 3.2 Non-rigid Adaptation of Vertebral Shapes

Finally, local shape geometry for each of the ADM's vertebrae is obtained by varying the weight parameters of the principal variations. We parameterize these potentials with a set  $\mathcal{C}$  of clique variables  $c$ , controlled by high-order potential  $V_c$  [8] which assigns a cost to a configuration of  $c$ . Each clique are assigned to weight vectors  $\omega_c$ . Hence the third term of (3) is described as a high-order functional:

$$\beta V(\mathbf{H}, \Delta, \Omega) = \beta \sum_{c \in \mathcal{C}} V_c(\mathbf{w}_c^0 + \omega_c) \quad (7)$$

where independent clique variables  $c$  are treated as a graph minimization problem. The prior term is represented by higher-order potentials of degree  $n$ , based on the eigenvalues of the  $L$  local vertebrae from our model  $\mathbf{S}$ . Our work is inspired from a mesh reconfiguration [9] where costs are associated to cliques  $c$  based on the positions of the morphed mesh vertices  $\mathbf{v}_i$ . A search is performed along the normal  $\mathbf{n}_i$  from  $\mathbf{v}_i$  to find the optimal compromise between boundary detection and the distance to the mean eigenvalue shape. We therefore penalize deformations which deviates from the local distribution.

One can integrate the global data and prior terms along with local shape terms parameterized as the higher-order cliques, by combining (4), (6) and (7):

$$E(\mathbf{S}^0, \mathcal{I}, \Delta, \Omega) = V(f_{\text{NW}}(\{y_1 + \delta_1, \dots, y_d + \delta_d\}, \mathcal{I})) + \alpha \sum_{i \in G} \sum_{j \in \mathcal{N}(i)} V_{ij}(y_i^0 + \delta_i, y_j^0 + \delta_j) + \beta \sum_{c \in \mathcal{C}} V_c(\mathbf{w}_c^0 + \omega_c). \quad (8)$$

### 3.3 Energy Minimization

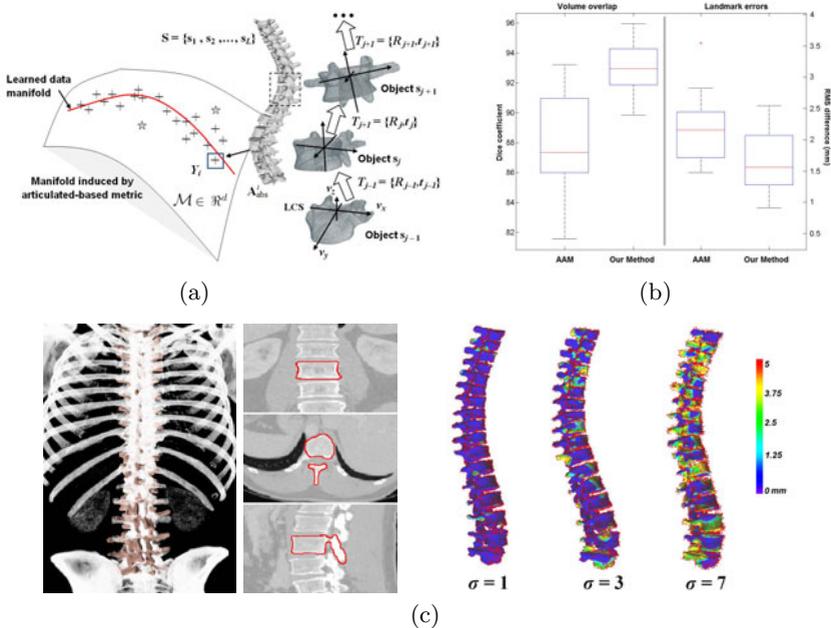
The optimization strategy of the resulting MRF (8) in the continuous domain is not a straightforward problem. The convexity of the solution domain is not guaranteed, while gradient-descent optimization approaches are prone to non-linearity and local minimums. We seek to assign the optimal labels  $\mathcal{L}^\Delta = \{l_1, \dots, l_d\}$  and  $\mathcal{L}^\Omega = \{l_1, \dots, l_n\}$  which are associated to the quantized space  $\Delta$  of displacements and local weight parameters  $\Omega$  respectively. We consider that displacing the coordinates of point  $y_i^0$  by  $\delta^{l_i}$  is equivalent to assigning label  $l_i$  to  $y_i^0$ . An incremental approach is adopted where in each iteration  $t$  we look for the set of labels that improves the current solution s.t.  $y_i^t = y_i^0 + \sum_t \delta^{l_i t}$ , which is a temporal minimization problem. Then (8) can be re-written as:

$$E^t(\mathcal{L}^\Delta, \mathcal{L}^\Omega) = V(f_{\text{NW}}(\{y_1^{t-1}, l_1^\Delta, \dots, y_d^{t-1}, l_d^\Delta\}, \mathcal{I})) + \alpha \sum_{i \in G} \sum_{j \in \mathcal{N}(i)} V_{ij}(y_i^{t-1}, y_j^{t-1}, l_i^\Delta, l_j^\Delta) + \beta \sum_{c \in \mathcal{C}} V_c(\mathbf{w}_c^{t-1}, l_c^\Omega). \quad (9)$$

We solve the minimization of the higher-order cliques in (9) by transforming them into quadratic functions [8]. We apply the FastPD method [10] which solves the problem by formulating the duality theory in linear programming.

### 4 Experiments and Results

In order to evaluate the performance, we considered modeling pathological spinal columns for CT inference. We used a database of  $N = 711$  spine models reconstructed in 3D and exhibiting different types of deformations relative to global and local shapes. For each spine, 6 landmarks on each of the 17 vertebrae composing the spinal column were used to extrapolate the inter-object transformation (Fig. 2(a)). Optimal neighborhood size was found at  $K = 10$ , while intrinsic dimensions was  $d = 7$  and  $n = 5$ , dictating the number of nodes in our global graph model.  $\alpha$  and  $\beta$  balance the contribution of the energy terms and set at 0.3 and 0.5 respectively. We tested the algorithm on a subset of 20 unseen 3D reconstructed cases from the database (modifying the modular data term to vertex correspondences) and 12 CT volumes. We quantitatively compared our method to an AAM modeling based on global PCA. Dice scores and root-mean-square (RMS) landmark distances show improvement of the proposed MRF approach via a non-linear shape analysis in Fig. 2(b). Successful examples from CT inferred data are shown in Fig. 2(c). We evaluated the performance based on the density of the input sample points, affecting the global shape inference starting when only 20% of points are available. Furthermore, added Gaussian distributed noise



**Fig. 2.** (a) Representation of inter-vertebral transformations. (b) Plots of Dice coefficients and landmark errors comparing our method to AAM. (c) Spine inferences of triangulated from CT with orthogonal views. Local shape distortions with significant noise level increase  $\sigma$  added to target points (error-coded for ground-truth distances).

to the target data introduces distortions when  $\sigma = 3$  (Fig. 2). These validations prove how our method elegantly encodes prior knowledge with image constraints in an MRF framework, and efficiently minimizes the energy term to converge towards an optimal solution. One drawback remains the computational time due to the inverse regression mapping and higher-order clique potential minimization.

## 5 Discussion and Future Work

Our main contribution consists in modeling complex, non-linear patterns of spine deformations in a Riemannian manifold. Point-based models are created from statistical knowledge in terms of global and local variations. To this end, an articulated distance metric based on intrinsic and orientation properties was proposed. The non-linear embedding is constructed in such a way to avoid creating shape distortions, as well as collisions between shapes. This is accomplished by constraints within the manifold parameters to restrict outlying values corrupting global shape appearance. The proposed framework can be extended in medical imaging applications to allow global/local shape modeling. While the method remains computationally expensive for the inverse map, inference of articulated models based on optimal clique decomposition can be beneficial to this end.

## References

1. Nain, D., Haker, S., Bobick, A., Tannenbaum, A.: Multiscale 3-D shape representation and segmentation using spherical wavelets. *T. Med. Imag.*, 598–618 (2007)
2. de Bruijne, M., Lund, M., Tanko, L., Pettersen, P., Nielsen, M.: Quantitative vertebral morphometry using neighbor-conditional shape models. In: Larsen, R., Nielsen, M., Sporrang, J. (eds.) *MICCAI 2006*. LNCS, vol. 4190, pp. 1–8. Springer, Heidelberg (2006)
3. Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., Ayache, N.: Articulated spine models for 3-D reconstruction from partial radiographic data. *T. Biomed. Eng.* 11, 2565–2574 (2008)
4. Klinder, T., Wolz, R., Lorenz, C., Franz, A., Ostermann, J.: Spine segmentation using articulated shape models. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I*. LNCS, vol. 5241, pp. 227–234. Springer, Heidelberg (2008)
5. Kadoury, S., Paragios, N.: Surface/volume-based articulated 3D spine inference through Markov Random Fields. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009*. LNCS, vol. 5762, pp. 92–99. Springer, Heidelberg (2009)
6. Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. *Science* 290, 2323–2326 (2000)
7. Davis, B., Fletcher, P., Bullitt, E., Joshi, S.: Population shape regression from random design data. In: *ICCV*, pp. 1–7 (2007)
8. Rother, C., Kohli, P., Feng, W., Jia, J.: Minimizing sparse higher order energy functions of discrete variables. In: *CVPR*, pp. 1382–1389 (2009)
9. Kaus, M., Pekar, V., Lorenz, C., Truyen, R., Lobregt, S., Weese, J.: Automated 3D PDM construction from segmented images using deformable models. *IEEE T. M. Imag.* 22, 1005–1013 (2003)
10. Komodakis, N., Tziritas, G., Paragios, N.: Performance vs computational efficiency for optimizing single and dynamic mrf: Setting the state of the art with primal dual strategies. *CVIU* 112(1), 14–29 (2008)