Shape Representation via Best Orthogonal Basis Selection

Ashraf Mohamed^{1,2} and Christos Davatzikos^{1,2}

¹ CISST NSF Engineering Research Center,
Department of Computer Science, Johns Hopkins University
<http://cisstweb.cs.jhu.edu/>

² Section for Biomedical Image Analysis, Department of Radiology,
University of Pennsylvania School of Medicine
<http://oasis.rad.upenn.edu/sbia/>
{ashraf,christos}@rad.upenn.edu

Abstract. We formulate the problem of finding a statistical representation of shape as a best basis selection problem in which the goal is to choose the basis for optimal shape representation from a very large library of bases. In this work, our emphasis is on applying this basis selection framework using the wavelet packets library to estimate the probability density function of a class of shapes from a limited number of training samples. Wavelet packets offer a large number of complete orthonormal bases which can be searched for the basis that optimally allows the analysis of shape details at different scales. The estimated statistical shape distribution is capable of generalizing to shape examples not encountered during training, while still being specific to the modeled class of shapes. Using contours from two-dimensional MRI images of the corpus callosum, we demonstrate the ability of this approach to approximate the probability distribution of the modeled shapes, even with a few training samples.

1 Introduction

The statistical study of anatomical shape variability in medical images allows for the morphometric characterization of differences between normal and diseased populations [1,2,3], the detection of shape changes due to growth or disease progression [4,5], the construction of models capable of automatically segmenting and labelling structures in images [6,7], and the estimation of missing shape features from observed ones [8].

A space of high dimensionality is typically needed to represent shapes in medical images faithfully. With the limited availability of training samples, the statistical shape modeling and analysis task becomes infeasible in the original high dimensional space. This problem has traditionally been approached by applying dimensionality reduction methods, such as principal component analysis (PCA), to yield a compact shape representation (shape descriptor) with a small number of parameters that can be estimated easily from the training samples. However, with the limited availability of training samples, PCA becomes sensitive

to noise and data outliers, and it allows only a small number of degrees of freedom for the shape descriptor, which makes it unable to generalize to valid shape instances not encountered during training.

In this work, we propose the use of the best orthogonal basis selection framework of Coifman and Wickerhauser [9] to decide a multi-resolution shape representation that is optimal according to an application-dependent criterion. Multi-resolution analysis offers a shape representation with a large number of degrees of freedom by analyzing shape at different scales independently [10]. For example, wavelet packets (WPs) – of which the discrete wavelet transform is a special case – offers a large library of complete orthonormal bases, each of which capable of linearly mapping shape onto a different multi-resolution feature space. This library can be searched for the optimal basis according to a criterion that depends on the shape analysis task at hand. We demonstrate the use of this basis selection framework for approximating the probability density function (pdf) of a class of shapes by minimizing the entropy of WP coefficients [11]. This yields wavelet coefficients that are as decorrelated as possible, and therefore, coupled with a Gaussian assumption, allows the expression of the pdf as the product of lower dimensional pdfs that can be estimated more accurately from the limited available training samples. Such estimates of shape pdfs are useful for computing the likelihood of shapes in classification problems, and can constitute a shape prior to be used in generative models for shape segmentation and estimation tasks.

The rest of the paper is organized in four sections. In Section 2, we explain our shape representation and review the use of PCA for shape modeling and analysis. In Section 3, we present the details of our approach with a brief overview of wavelet packets and best basis selection. In Section 4, using contours of two-dimensional (2D) corpora callosa extracted from MRI images, we demonstrate the ability of our approach to approximate the pdf of the modeled shapes even with a few training samples. In Section 5, we explain how our approach is readily extendable to 3D images and we present an outlook for future work.

2 Shape Descriptors Using PCA

A continuous shape instance in $\mathbb{R}^d$ may be represented by a map from a d_r dimensional reference space to the shape space, $\psi : \mathbb{R}^{d_r} \rightarrow \mathbb{R}^d$. For example, in case of 3D shapes, $d = d_r = 3$ with ψ defined for each shape instance as the deformation map that registers this instance to a reference template [3]. For the sake of simplicity, we will restrict our discussion to shapes represented by a closed 2D contour which therefore, may be defined by $\psi : [0, 1) \rightarrow \mathbb{R}^2$ ($d_r = 1, d = 2$), with values in $[0, 1)$ being the normalized distance along the contour.

Assume that n_s examples are available from $\mathcal{C}$, the class of shapes of interest. Also, assume that Procrustes alignment was performed on the shapes to remove effects of translation, rotation, and scaling [6]. A uniform sampling of the interval $[0, 1)$ with n_p points ($n_p = 2^{K_o}$, and K_o a positive integer) is achieved by piece-wise constant speed reparameterization of the original contours obtained via manual or automatic landmark selection. Let the coordi-

nates $\{(x_i, y_i)\}_{i=1}^{n_p}$, of the resulting points around the contour be concatenated into a vector $\mathbf{x} = [x_1, \dots, x_{n_p}, y_1, \dots, y_{n_p}]^T \in \mathbb{R}^n$ ($n = 2n_p$). Therefore, the set $\mathcal{T} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_s}\}$, may be regarded as n_s independent realizations of a random vector $\mathbf{X}$ with an unknown pdf, $f_{\mathbf{X}}$.

For anatomical shapes, $n_s \ll n$ which makes shape analysis impractical or even infeasible in the original n -dimensional space. However, if $\mathcal{C}$ has an intrinsic dimensionality $m \ll n$ (e.g., due to correlations between the shape point coordinates), the analysis may be performed in a low-dimensional subspace $\mathcal{Q}$ of $\mathbb{R}^n$, which we hope to include $\mathcal{C}$. To analyze data in $\mathcal{Q}$, a map $\mathcal{V} : \mathbb{R}^n \rightarrow \mathcal{Q}$ must be defined. In this work, we restrict our attention to linear maps which can be defined by an invertible real $n \times n$ matrix $\mathbf{A}$, whose columns are basis vectors spanning $\mathbb{R}^n$. Let $\mathbf{X} = \mathbf{A}\mathbf{V}$. The map $\mathcal{V}$ is decided by choosing $\mathbf{A}^*$ as the optimal basis matrix according to some criterion that measures the efficiency of the basis with respect to $\mathcal{T}$, followed by selecting a collection $\{\mathbf{a}_i^*\}_{i=1}^r$ from among its columns.

Applying PCA to $\mathcal{T}$, yields the matrix $\mathbf{A}_{pca}^*$ whose columns are the eigenvectors of the sample covariance matrix of $\mathbf{X}$. The matrix $\mathbf{A}_{pca}^*$ is the rotation satisfying a number of optimality criteria including attaining the minimum of $h_{\mathbf{X}}$, the entropy function of the coefficients $\mathbf{V} = \mathbf{A}^T \mathbf{X}$ defined by [12]:

$$h_{\mathbf{X}}(\mathbf{A}) = \sum_{i=1}^n \hat{\sigma}_i^2(\mathbf{A}) \log \hat{\sigma}_i^2(\mathbf{A}), \quad (1)$$

with $\hat{\sigma}_i^2(\mathbf{A}) = E[v_i^2] / \sum_{j=1}^n E[v_j^2]$, and v_i 's being components of $\mathbf{V}$. The maximum number of degrees of freedom (dof) of the shape descriptor $\mathbf{V}$ is $n_s - 1$ which is the number of eigenvectors of the sample covariance matrix with non-zero eigenvalues. These eigenvectors span the linear space in which the training samples lie. Therefore, if $m < n_s - 1$, $\mathbf{V}$ will be unable to represent legal shapes outside the linear subspace spanned by the training samples.

3 Methods

The abovementioned drawbacks of PCA may be addressed in the best basis selection (BBS) framework described in detail in [9,12]. In this context of shape analysis, the goal of BBS is to find the optimal representation of $\mathbf{X}$ in terms of the columns of $\mathbf{A}$ which are selected from the components of a library $\mathcal{L}$ of overcomplete bases ($|\mathcal{L}| > n$). The choice of the library $\mathcal{L}$ and the optimization criterion are task dependent. For example, for shape classification, $\mathcal{L}$ may be designed to test one or more hypothesis of local shape differences between two populations, while the optimization criterion would encourage the representation of shape in terms of the basis that provides the best clustering [13]. In general, $\mathbf{A}$ need not have n columns or constitute a complete basis. However, this will be the case in our work here.

We now direct our attention to the task of estimating $f_{\mathbf{X}}$ given $\mathcal{T}$. We assume that $\mathcal{L}$ is formed of orthonormal bases, such as the wavelet packets (WP)

library [11], briefly explained in Section 3.1. Any complete minimal basis from this library will correspond to a rotation matrix $\mathbf{A}_S$. Such a basis can be partitioned into N groups of basis vectors, with each group spanning a subspace $\mathcal{S}_l$, and the subspaces $\{\mathcal{S}_l\}_{l=1}^N$ mutually orthogonal. The partitioning is decided so that the subspaces $\{\mathcal{S}_l\}_{l=1}^N$ will correspond to shape features that can be analyzed independently of the others. One obvious, example would be localized shape details, and global shape characteristics. Defining $\mathbf{x}_p^l = \text{Proj}(\mathbf{x}, \mathcal{S}_l)$, we can write

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{l=1}^N f_{\mathbf{x}_p^l}(\mathbf{x}_p^l). \quad (2)$$

If each of the subspaces in $\{\mathcal{S}_l\}_{l=1}^N$, is of low dimensionality, then the independence assumption has effectively decomposed the pdf into the product of marginal pdfs that can be estimated more accurately from the limited samples. Moreover, it improves the generalization ability of the estimated pdf over that of PCA since the shape descriptor can have up to $n_s - 1$ dof from each subspace $\mathcal{S}_l$ compared to a total of $n_s - 1$ dof in PCA. On the other hand, high correlation between $\mathbf{x}_p^l$ implied by the training samples, may betray the invalidity of this assumption, and therefore suggests that the improved generalization ability of the estimated $f_{\mathbf{X}}$ comes at the expense of its low specificity (i.e., the estimated $f_{\mathbf{X}}(\mathbf{x})$ may include invalid shapes). Therefore, we select the optimal basis $\mathbf{A}_S$ according to an information cost that, coupled with a Gaussian assumption on $f_{\mathbf{X}}$, supports independence of shape projections on the subspaces $\{\mathcal{S}_l\}_{l=1}^N$. This criterion is explained in Section 3.2.

3.1 The Wavelet Packet Bases Library

Wavelet packets (WPs) form a family of orthonormal bases for discrete functions in $L^2(\mathbb{R})$ [9]. Dyadic multi-resolution analysis of a signal may be performed by defining wavelet analysis functions $\{W_k\}_{k=0}^\infty$ associated with an exact quadrature mirror filter pair, $h(i)$ and $g(i)$, as:

$$\begin{aligned} W_{2i}(z) &= \sqrt{2} \sum h(k) W_i(2z - k) \\ W_{2i+1}(z) &= \sqrt{2} \sum g(k) W_i(2z - k) \end{aligned} \quad (3)$$

The functions $W_0(z)$ and $W_1(z)$ correspond to the mother wavelet and the scaling function respectively.

A wavelet packet basis (WPB) of $L^2(\mathbb{R})$ is any orthonormal basis selected from the functions $W_{k,i,j} \equiv 2^{k/2} W_i(2^k t - j)$ [9]. For an analyzed signal of length 2^{K_o} , each $W_{k,i,j}$ is identified by the indices

- $k = 0, \dots, K$: the scale parameter, where K is the maximum scale ($K \leq K_o$),
- $i = 0, \dots, 2^k - 1$: the frequency index at scale k , and
- $j = 0, \dots, 2^{K_o-k} - 1$: the translation index at scale k .

The library of WPs may be organized into a binary tree with each level representing the projection of the signal of the parent node onto two subspaces corresponding to its low frequency component and high frequency details. If each wavelet packet $W_{k,i,j}$ is represented by a $2^{K_o} \times 1$ vector, we can define

$$\mathbf{W}_{k,i} = [W_{k,i,0} | \cdots | W_{k,i,2^{K_o}-k-1}] \quad (4)$$

to be the collection of WPs for all translations at node (k, i) of the tree. If a collection $\{\mathbf{W}_{k,i}\}$ corresponds to a disjoint cover of the frequency domain, then it forms an orthonormal basis [9]. The discrete wavelet transform constitutes one of these basis. Every such basis will correspond to the leaves of a full binary wavelet packet tree (WPT). There are 2^{2^K} such bases for a binary tree of K levels. This flexibility can be exploited to select the most efficient representation of a signal according to some information cost.

3.2 Best Basis Selection (BBS) for Estimating 2D Shape pdf

The shape vector $\mathbf{X}$, may be decomposed into two signals written as column vectors $\mathbf{X}^1 = [x_1, \dots, x_{n_p}]^T$ and $\mathbf{X}^2 = [y_1, \dots, y_{n_p}]^T$, with $\mathbf{X} = [\mathbf{X}^1 \mathbf{X}^2]^T$. We elect to use the same WPB for both signals. Any set $\{\mathbf{W}_{k,i}\}$ forming a WPB of $\mathbb{R}^{n_p}$ may be written in the form an $n_p \times n_p$ basis matrix $\mathbf{B}$. The matrix $\mathbf{A}_{wpt} = \text{diag}(\mathbf{B}, \mathbf{B})$ forms a basis for $\mathbb{R}^n$, and the shape descriptor is $\mathbf{V} = \mathbf{A}_{wpt}^T \mathbf{X} = [\mathbf{V}^1 \mathbf{V}^2]^T$ with $\mathbf{V}^1 = \mathbf{B}^T \mathbf{X}^1$ and $\mathbf{V}^2 = \mathbf{B}^T \mathbf{X}^2$.

The goal of the BBS algorithm is to find the matrix $\mathbf{A}_{wpt}^*$ attaining the minimum of an information cost functional $\mathcal{M}_{\mathbf{X}}(\mathbf{A})$ that measures the efficiency of the basis $\mathbf{A}_{wpt}$ in representing $\mathbf{X}$. Defining the function $\mathcal{M}$, as a map from real sequences $\{v_i\}$ to $\mathbb{R}$, we can now define $\mathcal{M}_{\mathbf{X}}(\mathbf{A}) \equiv \mathcal{M}(\{v_i\})$, with v_i 's the components of $\mathbf{V} = \mathbf{A}^T \mathbf{X}$. Coifman and Wickerhauser [9] proposed a fast BBS algorithm suitable for searching the WP library when $\mathcal{M}$ is additive, (i.e., $\mathcal{M}(0) = 0$ and $\mathcal{M}(\{v_i\}) = \sum_i \mathcal{M}(v_i)$). Here, we use $\mathcal{M}_{\mathbf{X}}(\mathbf{A}_{wpt}) = h_{\mathbf{X}^1}(\mathbf{B}) + h_{\mathbf{X}^2}(\mathbf{B})$, where $h_{\mathbf{X}^1}(\mathbf{B})$ and $h_{\mathbf{X}^2}(\mathbf{B})$ are defined analogous to equation (1). It can be shown that since $h_{\mathbf{X}^l}$ is an additive information cost for $l = 1, 2$ [11], so will $\mathcal{M}$ be. This choice of $\mathcal{M}$ will provide the optimal WPB, $\mathbf{B}^*$, and therefore $\mathbf{A}_{wpt}^*$, that produces $\mathbf{V}^1$ and $\mathbf{V}^2$, each composed of a set of coefficients that are as decorrelated as possible. This basis is called the Joint Best Basis (JBB) [11].

The shape descriptor $\mathbf{V}$ may be decomposed into a collection $\{\mathbf{v}_{k,i}\}$ of WP coefficients with $\mathbf{v}_{k,i} \equiv [(\mathbf{W}_{k,i}^{*T} \mathbf{x}^1)^T (\mathbf{W}_{k,i}^{*T} \mathbf{x}^2)^T]^T$ for the set $\{\mathbf{W}_{k,i}^*\}$ which forms the optimal basis $\mathbf{B}^*$. Here, we assume that each $\mathbf{v}_{k,i}$ is independent of $\mathbf{v}_{u,v}$ for all $(k, i) \neq (u, v)$. With this choice, it can be seen that each $\mathbf{W}_{k,i}^{*T}$ describes one of the subspaces S_l mentioned above. In the work presented here, each coefficients vector $\mathbf{v}_{k,i}$ is assumed to follow a Gaussian distribution and PCA is applied to estimate this distribution [13].

4 Experimental Results

We compare the estimated shape pdfs generated by PCA basis (PCAB), JBB, and the approach in [10] which will be referred to as WT basis (WTB). We

Fig. 1. Example MRI images with manually outlined corpus callosum contours.

use a dataset composed of 100 manually-outlined corpora callosa contours in midsagittal MR images of normal subjects. Landmarks were manually selected by an expert and the contours were then reparametrized by piece-wise constant speed parameterization. We used $n_p = 512$, and therefore $K_o = 9$. Example contours from the dataset are shown in Figure 1. More details about the datasets can be found in [10]. For wavelet analysis, we used the Daubechies wavelet with 6 vanishing moments and $K = 5$ (5 levels of wavelet decomposition).

Fig. 2. Average generalization (left) and specificity (right) errors for the estimated $f_{\mathbf{x}}$ by JBB, WTB, and PCAB.

Ideally, an estimated $f_{\mathbf{x}}$ may be evaluated by comparing it with the actual pdf responsible for generating the analyzed shapes. For naturally occurring shapes, the latter, is however, not known. Therefore, we adopt the measures of generalization and specificity used in [14]. We do not address the issue of compactness here since our goal is to increase the number of degrees of freedom over that of PCA. PCA will always produce a more compact representation than WTB or JBB for any percentage of the total variance observed in the training samples. Constructing a more compact shape descriptor may be achieved via wavelet shrinkage or the approach in [13] and is subject to future investigations.

We introduce the maximum absolute distance (MAD) across all corresponding points in two shapes as a measure of the similarity between them. Cross validation via the leave-one-out method is used for measuring the generalization

Fig. 3. (a) The JBB tree compared to the WTB tree for 25 training samples. Circles denote the terminal nodes of the WTB. The length of each branch in the tree indicates the amount of entropy reduction achieved by splitting a node into its two children. (b) Zoom-in to the reconstruction of a corpus callosum contour using PCAB and the JBB tree in (a).

ability of an the estimated $f_{\mathbf{X}}$. The generalization error for the left-out sample is defined as the MAD between this sample and its best possible reconstruction using the model. Since each of the tested models is composed orthonormal vectors, the best-possible reconstruction of a shape is found by projecting it on the model's basis vectors, followed by restricting the coefficients of the basis to be within three standard deviations from the mean as estimated from the training samples. The average generalization error of a model can then be computed over the complete set of trials which involve leaving-out a different sample each time.

To evaluate the specificity, a shape is randomly generated according to the estimated pdf, and MAD is computed between this shape and all available samples, whether used for estimating $f_{\mathbf{X}}$ or not. The minimum MAD is recorded as the specificity error. The average specificity error is then computed for 1000 randomly generated shapes.

The average generalization and specificity errors for varying number of training samples are shown in Figure 2. JBB and WTB estimates of $f_{\mathbf{X}}$ have better generalization ability than that of PCAB. On the other hand, the specificity error is higher for both JBB and WTB compared to PCAB. The increase in the specificity error of JBB over that of PCA is maximum of 9% (with 10 training samples). JBB has better specificity than WTB for 50 training samples or less.

In Figure 3.a, the WPT of JBB for 25 training samples is compared to that of the WTB. The structure of the two trees are similar, although not identical. Some bands in the WTB are split into their two children in the JBB because the coefficients of the children offer a lower entropy. JBB has a total of 10 nodes giving rise to a maximum of 240 dof while WTB has 6 nodes, which gives rise to a maximum of 144 dof. On the other hand PCAB has a maximum of 24 dof.

To qualitatively evaluate the generalization ability of the JBB prior in this case, in Figure 3.b reconstructions of a corpus callosum contour (not used in training) is shown for PCAB and JBB. The JBB reconstruction follows the original contour closely, while the PCAB contour shows some significant differences.

5 Discussion and Future Work

We presented the problem of finding a shape descriptor as a BBS problem in which the goal is to select the optimal basis for shape representation from a very large library of bases. Within this framework of BBS, we demonstrated the use of the WP library for estimating the pdf of shapes of 2D corpus callosa contours. The independence assumption between the wavelet coefficients at different nodes of the selected WPT allows the expression the pdf as a product of lower dimensional pdfs which can easily be estimated from the training samples. This assumption reveals similarities between our approach and that of Staib and Duncan [15], although ours enjoys the advantages of wavelet analysis over Fourier analysis. The assumption is also equivalent to imposing a block diagonal prior on the structure on the covariance matrix of $\mathbf{V} = \mathbf{A}_S^T \mathbf{X}$ [10]. This observation connects our approach to that in [7] and the references therein, which have the effect of imposing some form of a structure on the covariance matrix of $\mathbf{X}$. A key difference is that transforming shapes using $\mathbf{A}_S$ allows the shape description in a basis for which prior knowledge about shape (e.g., independence between local and global shape) has direct a implication on the covariance structure that is not a result of an ad hoc physical or general smoothness model.

As pointed out in the body of this work, with the use of other orthogonal basis libraries and selection criteria, many shape analysis applications can benefit from the approach presented here. With the availability of deformation fields mapping shape to a standard template, the extension of this work to 3D and 4D shapes is straight forward. Usually, such deformation fields are readily available as a product of an image warping procedure, which is essential for establishing automatic correspondence between shapes in 3D and 4D images. The application of this approach for 3D shapes is currently underway. Also, we plan to investigate other approaches for the grouping of basis vectors rather than using the nodes of the WPT, and approaches to achieve more compact descriptors.

Acknowledgments. The authors would like to thank Xiaodong Tao for providing the reparametrized corpus callosum data. WavPack library was used for programing and creating some figures in this work. This work was supported in part by the National Science Foundation under Engineering Research Center grant EEC9731478, and by the National Institutes of Health grant R01NS42645.

References

1. P. Golland, W. E. L. Grimson, M. E. Shenton, and R. Kikinis. Deformation analysis for shape based classification. In *Proc. of IPMI'2001*, LNCS 2082, pages 517–530, 2001.
2. P. Yushkevich, S. Joshi, S. M. Pizer, J. G. Csernansky, and L. E. Wang. Feature selection for shape-based classification of biological objects. In *Proc. of IPMI'2003*, LNCS 2732, pages 114–125, 2003.

3. J. G. Csernansky, S. Joshi, L. Wang, J. W. Haller, M. Gado, J. P. Miller, U. Grenander, and M. I. Miller. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. In *Proc. Natl. Acad. Sci.*, 95:11406–11411, Sept. 1998.
4. P. R. Andersen, F. L. Bookstein, K. Conradsen, B. K. Ersbøll, J. L. Marsh, and S. Kreiborg. Surface-bounded growth modeling applied to human mandibles. *IEEE Trans. Med. Imag.*, 19(11):1053–1063, Nov. 2000.
5. C. Davatzikos, A. Genc, D. Xu, and S. M. Resnik. Voxel-based morphometry using the ravens maps: Methods and validation using simulation of longitudinal atrophy. *Neuroimage*, 14:1361–1369, Dec. 2001.
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(1):38–59, 1995.
7. Y. Wang and L. H. Staib. Boundary finding with prior shape and smoothness models. *IEEE Trans. PAMI*, 22(7):738–743, July 2000.
8. A. Mohamed, S. K. Kyriacou, and C. Davatzikos. A statistical approach for estimating brain tumor induced deformation. *Proc. of the IEEE MMBIA Workshop*, pages 52–59, December 2001.
9. R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. on Inf. Theory*, 38(2), March 1992.
10. C. Davatzikos, X. Tao, and D. Shen. Hierarchical active shape models using the wavelet transform. *IEEE Trans. on Med. Imag.*, 22(3):414–423, March 2003.
11. M. V. Wickerhauser. Fast approximate factor analysis. In *Curves and Surfaces in Computer Vision and Graphics II*, In *SPIE Proc.*, 1610:23–32, Boston, 1991.
12. N. Saito. Image approximation and modeling via least statistically dependent basis. *Pattern Recognition*, 34:1765–1784, 2001.
13. F. G. Meyer and J. Chinrungrueng. Analysis of event-related FMRI data using best clustering basis. *IEEE Trans. on Med. Imag.*, 22(8):933–939, Aug. 2003.
14. M. A. Styner, K. T. Rajamani, L. R. Nolte, G. Zsemlye, Gábor Székely, C. J. Taylor, and R. H. Davies. Evaluation of 3d correspondence methods for model building. In *Proceedings of IPMI'2003*, 2003.
15. L. H. Staib and J. S. Duncan. Boundary finding with parametrically deformable models. *IEEE Trans. on PAMI*, 14(11):1061–1075, Nov. 1992.