Scale-Space on Image Profiles
about an Object Boundary

Sean Ho⋆ and Guido Gerig

Department of Computer Science
University of North Carolina, Chapel Hill, NC 27599, USA
`seanho@cs.unc.edu`

Abstract. Traditionally, image blurring by diffusion is done in Euclidean space, in an image-based coordinate system. This will blur edges at object boundaries, making segmentation difficult. Geometry-driven diffusion [1] uses a geometric model to steer the blurring, so as to blur along the boundary (to overcome noise) but edge-detect across the object boundary. In this paper, we present a scale-space on image profiles taken about the object boundary, in an object-intrinsic coordinate system. The profiles are sampled from the image in the fashion of Active Shape Models [2], and a scale-space is constructed on the profiles, where diffusion is run only in directions tangent to the boundary. Features from the scale-space are then used to build a statistical model of the image structure about the boundary, trained on a population of images with corresponding geometric models. This statistical image match model can be used in an image segmentation framework. Results are shown in 2D on synthetic and real-world objects; the methods can also be extended to 3D.

1 Introduction

Features in images exist at scale, hence scale-spaces of images are useful for characterizing them. In image segmentation, we wish to localize the boundary of an object within the image. The traditional linear scale space can blur the boundary, complicating segmentation. Early on, Canny's [3] edge-detection work motivated a scheme which blurs along the boundary, but edge-detects across the boundary. Variable-conductance diffusion [4] and geometry-driven diffusion [1] aim to achieve this, by using the geometry of the image intensity field to orient and modulate the diffusion locally.

In deformable model-based segmentation, the deformable model has geometry of its own. In this paper, we develop an implementation of a geometry-driven scale-space which uses the geometry of a deformable model instead of the geometry of the image intensity field to drive the diffusion. This scale-space in an object-intrinsic coordinate system is then used to construct a statistical model of image structure about the object boundary, as trained on a population of segmented images. The statistical model forms the image-match likelihood, which

---

⋆ Supported by NIH-NCI P01 CA47982.

Fig. 1. Sagittal slice of the hippocampus in MRI, showing complex boundary appearance. The hippocampus is outlined in red. The tail of the hippocampus is to the right; the head is to the left. The hippocampus is a 3D object; this image only shows a cross-section.

when coupled with a prior on the shape of the deformable model, forms a complete segmentation framework.

Objects in images, particularly medical images, often exhibit complex boundary profiles, which can not be represented by simple step edges. Hence the image gradient magnitude map alone is not sufficient to drive segmentation of these objects. Frequently an object in a medical image will have different image profiles at different locations around the boundary. For instance, the hippocampus in T1-weighted MRI shows a sharp step edge at the tail from grey inside the hippocampus to black outside (Figure 1). However, parts of the bottom of the hippocampus might be characterized by a bar-like edge, and at the head of the hippocampus, the transition zone to the adjacent amygdala is not always visible, resulting in no edge at all in the profile at that location. Successful segmentation of these real-world objects in real-world images that have noise and large variability necessitates the development of an image model which can represent these complex boundary profiles; one that is local, statistical, and multiscale.

2   Related Work

2.1   Geometry-Driven Diffusion

Viewing an image as an embedding map between Riemannian manifolds, e.g. viewing the intensity field as a height field, provides the framework to apply principles from Riemannian differential geometry to image analysis. Much work has been done on geometry-driven diffusion [1], where the geometry which drives the diffusion is taken from the image itself. Mean curvature flow and anisotropic diffusion [4] are well-known examples. The Beltrami flow [5] smooths the image while preserving edges by using the Laplace-Beltrami diffusion operator on the manifold given by the intensity field. The Laplace-Beltrami operator is the generalization of the Laplacian from Euclidean space to arbitrary manifolds, incorporating the Riemannian metric tensors of the manifold.

Sapiro et al [6] apply Laplace-Beltrami diffusion to directional data defined on a flat planar domain, although the theory they develop is applicable to maps between arbitrary manifolds.

Chung et al [7] use the Laplace-Beltrami operator to perform diffusion on a manifold given by a triangulation of the brain cortical surface, to perform diffusion smoothing of a scalar field (e.g. mean curvature) defined on the manifold of the cortex. With this method, the geometry which drives the diffusion is no longer an image intensity field, but rather externally provided geometry, in the form of a triangle mesh outlining the cortical surface. The "image" is a map from the manifold of the cortical surface to the 1D range of mean curvatures; i.e. a texture map on a triangulated mesh.

2.2   Profile Model

Cootes and Taylor's seminal Active Shape Model work [2] samples the image along 1D profiles around boundary points, normal to the boundary, using correspondence given by the Point Distribution Model of geometry. Profiles are extracted at corresponding locations in a population of segmented training images. At each location along the boundary, a probability distribution is trained on the image-derived features in the 1D profile, and these probability distributions form the image-match model for segmentation.

The "hedgehog" model in the 3D spherical harmonic segmentation framework [8] can be seen as a variant of ASMs, and uses a training population linked with correspondence from the geometric model. It can be extended with a coarse-to-fine sampling of the object boundary in order to improve robustness in the face of image noise and jitter in correspondence.

2.3   Active Appearance Models

Perhaps the most well-developed work on modeling intensity variation in objects is the Active Appearance Model [9], which uses the point correspondences given by an ASM to warp images into a common coordinate system, within a region of interest across the whole object, given by a triangulation of the PDM. A global Principal Components Analysis is performed on the intensities across the whole object (the size of the feature space is the number of pixels in the region of interest). The use of a global PCA is particularly well-suited for capturing global illumination changes, for instance in Cootes and Taylor's face recognition applications.

It should be noted that both ASMs and AAMs have multiresolution extensions, which help to speed the optimization process of segmentation, and avoid local minima. However, the image pyramids used in the multiresolution extensions are created using isotropic diffusion in the original Euclidean image space, hence the object boundary will also be smoothed.

3   Method: Profile Scale-Space Model

We propose sampling intensities around the object boundary in an intrinsic coordinate system derived from the geometric model, which preserves the local
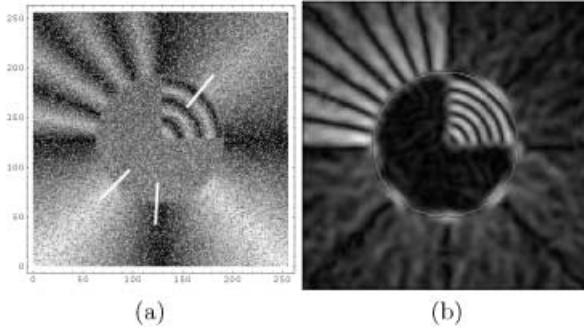
Fig. 2. (a) A sample from the toy example training population. (b) The corresponding gradient magnitude image at $\sigma = 4$ pixels, for reference. The contour from which profiles will be taken is shown in red. Each of the 30 images in the population have different additive noise, as well as multiplicative shading (linear ramp) in different directions.

orientation relative to the object boundary: i.e. which direction is across (normal to) the boundary and which directions are along (tangential to) the boundary. We want to look for edges across the boundary, but smooth away noise along the boundary. To accomplish this, we construct a scale-space in a non-Euclidean object-intrinsic coordinate system which preserves the along-boundary versus across-boundary distinction.

3.1   Toy Example

Throughout the discussion of our method, a toy example population of images will be used for ease of explanation, where the object shape is very simple (a circle in 2D), but the greyscale intensities are complex and have variability. Section 4 shows some results on real-world objects with shape variability. We use 2D Fourier coefficients [10] as a parameterized shape representation; the parameter space is the unit circle. The toy example images all have the same underlying structure, with sine waves of varying frequency at different sectors of the boundary, and boundary profiles more complex than a simple step edge. In addition, each image in the population has some global multiplicative shading in a different direction, and some uniform additive noise. The images are 256x256 pixels, and 30 sample images were created.

Figure 2 shows a sample from this toy population, as well as its gradient magnitude image at scale $\sigma = 4$ pixels. Note that a simple gradient magnitude search would not yield an appropriate image match model in this case, e.g. in the top right quadrant of the circle.

To illustrate the different edge structure at different locations around the boundary, profiles at three locations are shown in Figure 3. The locations from which the profiles are taken are shown overlaid on Figure 2(a) in yellow. The profile from the lower left quadrant of the circle shows a descending step edge
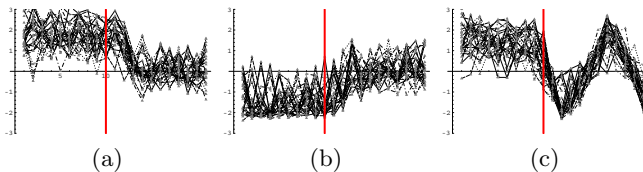
(a)   (b)   (c)

Fig. 3. Profiles from the training population, at the three boundary locations indicated in yellow in Figure 2 (counter-clockwise from lower left), showing: (a) negative step edge, (b) positive step edge, and (c) complex line-like edge. The x-axis indicates distance normal to the boundary (left is outside, right is inside the object). Object boundary is shown as red vertical line. The y-axis is image intensity.

from outside to inside the boundary; the profile at the bottom of the circle shows an ascending step edge, and the profile at the top right of the circle shows a complex structure with a line-like edge at the boundary and additional step edges further inside the object. Also shown in Figure 3 is the large variability in the training population, such that the step edge at places is overwhelmed by the image noise and variability.

3.2   Scale-Space in Object-Intrinsic Coordinates

Our object shape representation (2D Fourier harmonics) provides a mapping from a standardized parameter space to the object boundary in each image in the population. In the case of our toy example, the unit circle parameter space always maps to the same circle in each of the training images, but in a real-world population, the object shapes will differ within the training population. If we were to build a linear scale space in the original (x, y) image coordinates, edge structure across the boundary would also be blurred away. In the spirit of geometry-driven diffusion, we wish to blur along the boundary (to deal with noise), but detect edges across the boundary.

An elegant solution for diffusion along the object boundary is to map the blurring kernel from the parameter space onto the object. The image intensities sampled at the object boundary form a texture map on the surface of the object. The Laplace-Beltrami operator can then be used to blur this texture map. Extending the blurring to a collar of thickness about the boundary can be done by using the original object boundary to define a family of surfaces at different distances away from the object, e.g. a wavefront propagation using the distance transform on the boundary. Our implementation is an approximation to the Laplace-Beltrami diffusion.

In our implementation, we sample the image along 1D profiles normal to the boundary, unrolling the profiles into a flat rectangular grid. The "across-boundary" and "along-boundary" directions then become the x and y axes of the unrolled profile image, and we can construct a standard linear scale-space in one direction of the unrolled profiles, leaving the across-boundary direction unfiltered. The convolution in the along-boundary direction is cyclic, since the
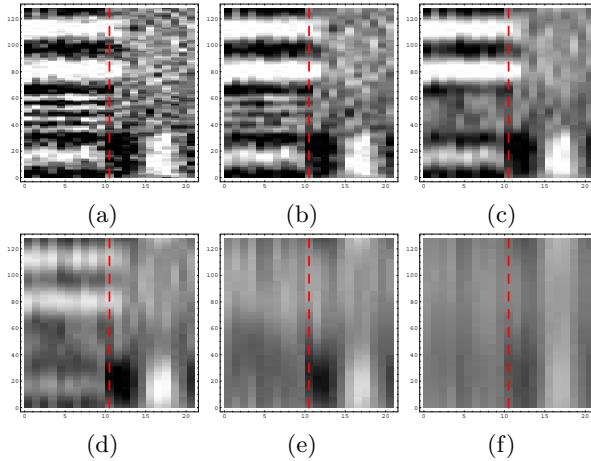
Fig. 4. The Gaussian scale-space constructed from one case in the toy example training population. Levels of the scale-space are (a)–(f), fine-scale to coarse-scale. Each row within a plot represents a boundary profile at that scale. The object boundary runs vertically through the center of each plot; left is exterior of the object and right is interior.

parameter space is the unit circle. We sample scale-space at $\sigma = 2, 4, 8, 16, 32, 64$. As a global brightness/contrast normalization pre-processing step, the profile images are scaled to have zero mean and unit variance, before the scale-space is constructed.

We build the Laplacian scale-space ($\partial/\partial\sigma$) on the profile image, representing each level of the scale-space as a residual from the next coarser level (Figure 5). This sets up a Markov chain in scale-space, as will be discussed in Section 3.3. One could also calculate local differences in space between neighboring profiles ($\partial/\partial\sigma \cdot \partial/\partial u$) to set up slightly different Markov neighborhood relations. So that the features are complete (we can reconstruct the original profiles from the features), we leave the coarsest level of the Gaussian scale-space (Figure 4(f)) alone in the Laplacian scale-space.

3.3   Local Statistical Model

The Laplacian scale-space is constructed on profiles taken from all images in the training population. We treat each profile at every location and scale level in the Laplacian scale-space as an independent feature vector, with its own PCA and (assumed Gaussian) distribution. Let $I(\sigma, u, \tau, i)$ be the normalized image intensity (a scalar), where:

- σ is the scale level in the Laplacian scale-space,
- u is the location of the profile around the boundary,
- τ is the position along the profile (in/out from the boundary),
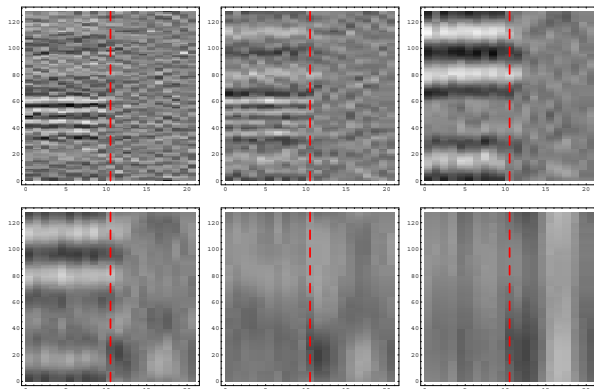- i is the index of the image in the training population.

Fig. 5. The corresponding Laplacian scale-space. Levels of scale-space are shown fine-to-coarse, as in Figure 4. The coarsest level is duplicated from the Gaussian scale-space.

With our toy example, each profile is 13 samples long, there are six scale levels, with 128 profiles around the boundary, and 30 images in the training population. The 13 samples along each profile are spaced in units of two pixels. At each combination of (σ, u) in the scale-space, we have a 13-dimensional feature space and a cluster of 30 points in it.

The statistical distributions are local to each scale and location, so that we can model local variability, such as the high-frequency sine waves in the northwest quadrant of our toy example. Because we use the Laplacian scale-space, the local distributions are tied together into one joint distribution through a Markov model, where we model only how each scale level differs from the next coarser scale level. The joint distribution over the whole profile image can be complex and non-Gaussian (Gibbs distribution), but each local distribution can be assumed to be Gaussian and hence easy to model. Non-parametric estimation of the local distributions (e.g. Parzen windowing), as in [11], could also be done instead of the Gaussian estimation.

By treating each profile in the scale-space as a feature vector, we avoid a model which is tied to a particular type of edge (step edge, bar-like edge, etc.), and this approach is different from feature selection from an n-jet [12] . The training population "tells us" what kind of edge to look for locally, including expected variability.

Figure 6 shows the mean profiles at all locations and scale levels in the scale-space. What is not shown is the local covariance structure which is also part of the statistical model. For instance, at the very coarse levels of the scale-space, the noise has been smoothed away, so that even though the edge structure is very weak, the variability across the training population is very small. At the finest scale level, some locations show a clear edge structure in the mean profile, however the variability due to noise at that scale overwhelms the step edge; compare with the scale-space of a single case in Figure 5.
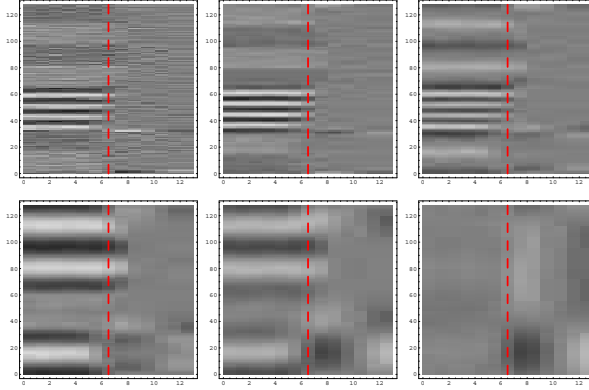
Fig. 6. Training population mean, Laplacian scale-space. Levels of the scale-space are shown fine-to-coarse, as in Figure 4.

Fig. 7. Midsagittal MRI slice of the brain, showing a corpus callosum from the training dataset. Overlaid in red is the manual segmentation, represented as a 2D Fourier contour.

4   Application to the Corpus Callosum

We have built a statistical profile scale-space model on 71 segmented 2D corpus callosum images, for use in a segmentation framework with a shape representation of 2D Fourier harmonics [13]. Each of the 71 corpora callosa have been hand-segmented by experts, and the Fourier harmonics constructed. An example from the training population is shown in Figure 7, together with its Fourier boundary.

Profiles were extracted from the images (128 profiles around the boundary), and the Laplacian scale-space built for each set of profiles. The large scale edge
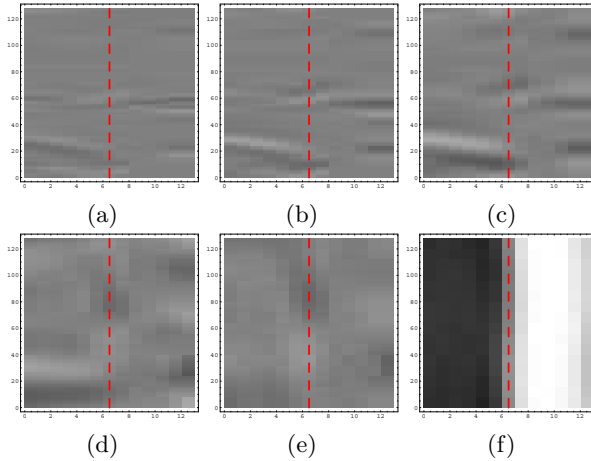
Fig. 8. Population mean of profiles around the corpus callosum; Laplacian scale-space. Levels of scale-space are shown fine-to-coarse, as in Figure 4.

structure of the corpus callosum can indeed be represented by a single step edge, which is in the very top level of the Gaussian scale-space, as shown in Figure 8(f).

Figure 8 shows the mean profiles of the Laplacian scale-space. As before there is also covariance structure behind each profile at each scale level that is not shown. The lower levels of the Laplacian scale-space show the local variability where the local edge structure differs from a simple step edge. The bright area of the fornix can be seen in levels 3 and 4 of the Laplacian scale-space, near the bottom of the graphs on the left hand side of each graph (exterior to the boundary). The fornix is a thin white-matter structure extending from the lower edge of the corpus callosum and not considered part of the corpus callosum; it can be seen in the MRI slice in Figure 7. Where the fornix joins the corpus callosum, there is no greylevel step edge at the boundary. The fornix is represented in the fine scales of the Laplacian pyramid as an *inverse* step edge (white outside, grey inside), which is the local residual needed to "cancel out" the global-scale step edge (dark outside, grey inside) shown in Figure 8(f).

5   Discussion / Conclusion

5.1   Application to Analysis and Discrimination

Our method constructs a statistical scale-space model of image profiles in a collar about the boundary of a deformable shape model placed in the image. Given a segmented training population, where each image in the population has a corresponding deformable shape model placed in it, our profile scale-space model allows one to understand how the image can be expected to vary along the object boundary. For example, the profile scale-space model built on the disk toy example clearly shows the sine waves of different frequency that are expected

at different locations about the boundary. As mentioned in the introduction, the hippocampus exhibits different edge profiles at different parts of the boundary. Our profile scale-space model can be used to quantify these different edge profiles locally, including natural variability in the population.

Another application of the profile scale-space model is in discrimination. Discrimination is often done on shape features to understand differences between subpopulations, e.g. a diseased patient group vs. a control group. Using local profile scale-space features, one could, instead of performing PCA, use discriminant analyses like Independent Component Analysis [14] to understand how the patient group differs from the control group in image features. For example, a clinical hypothesis might be that one portion of the hippocampus darkens slightly relative to the rest of the hippocampus in diseased patients. This (made-up) hypothesis is an example of what can be tested using discriminant analysis on the profile scale-space model.

In both analysis and discrmination, the profile scale-space model looks at image features after the shape variability has been factored away, by use of the correspondence given by the set of deformed shape models. As mentioned in Section 2.3, this is similar in spirit to Active Appearance Models [9]. PCA and discriminant analysis are often performed on shape features; similar analysis can be done on image features in a shape-normalized space. The image feature analysis is in complement to the shape analysis.

5.2   Application to Segmentation

Bayesian image segmentation by deformable models optimizes two competing terms, the image-match likelihood and the shape prior (shape typicality measure). The statistical model derived from the profile scale-space provides the image-match likelihood. Integrated with an appropriate shape prior and an optimization strategy, this can provide a complete segmentation framework. The deformable model is initialized in the target image, and the image profiles are sampled relative to the deformable model. A scale-space is built on the profiles, and features derived from the scale-space are compared to the scale-space features to the statistical model to get a goodness-of-fit of the deformable model into the target image. In addition to the global goodness-of-fit, local feature selection from the scale-space can allow a local confidence measure in the goodness-of-fit to be constructed. This goodness-of-fit is balanced against the shape prior, and the contour is deformed to optimize the combination. We have such a segmentation framework in 2D using Fourier harmonic shape descriptors; future work could extend this to 3D using analogous spherical harmonic shape descriptors.

5.3   Extension to 3D/nD Images

Geometry-driven diffusion literature has described generalized images as embedding maps between two manifolds. Chung et al [7] blur what amounts to a 2D texture map on a triangle mesh in 3D. If we let M be the triangle mesh (a 2D manifold embedded in 3D), then their image can be seen as a function

$I : M \rightarrow \mathbb{R}^+$, where $\mathbb{R}^+$ represents the positive reals, the range of intensity values in the image. Alternately, we can think of the image as a manifold embedded in the space $M \times \mathbb{R}^+$.

Now, to extend this to a collar of thickness about an object boundary (a volume of the image), the domain of the image can be extended to a 3D space about the boundary, instead of just the 2D manifold M. If we use the formalism of image profiles normal to the boundary, we can think of the image as a function $I : M \times [-1, 1] \rightarrow \mathbb{R}^+$, where $[-1, 1]$ indicates normalized distance from the object boundary M. Hence, for a fixed $d \in [-1, 1]$, the domain $\{(u, d) : u \in M\}$ represents a "shell" of distance d from the original boundary, according to a distance transform. For a fixed $u \in M$ on the surface, the set $\{I(u, d) : d \in [-1, 1]\}$ is a single image profile across the boundary at that point. Diffusion using the Laplace-Beltrami operator can then be run on each shell, using the metric tensor of the manifold M, to construct a scale-space. A non-Euclidean distance could certainly be used for d, for instance the intrinsic figural coordinate system in M-reps [15] [16].

5.4   Conclusion

We present a statistically trained scale-space model on image profiles around the boundary of an object embedded within images. There is a sampling of the image in a coordinate system relative to the object boundary (image profiles). A scale-space is constructed on these profiles which preserves the across-boundary features, but looks at along-boundary features in a multiscale fashion. The process is done for a population of images, with correspondence given by the shape models which have been deformed to fit each image. A statistical model is then built on the profile scale-space, which can then be incorporated into a full Bayesian segmentation framework.

References

1. Bart M. ter Haar Romeny, Ed., *Geometry-Driven Diffusion in Computer Vision*, Computational Imaging and Vision. Kluwer Academic Publisher, 1994.
2. Timothy F. Cootes, A. Hill, Christopher J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," in *IPMI*, 1993, pp. 33–47.
3. J.F. Canny, "A computational approach to edge detection," *IEEE Trans on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 8, no. 6, pp. 679–697, November 1986.
4. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 12, no. 7, pp. 629–639, July 1990.
5. N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 310–318, 1998.
6. Bei Tang, Guillermo Sapiro, and Vicent Caselles, "Diffusion of general data on non-flat manifolds via harmonic maps theory: The direction diffusion case," *International Journal of Computer Vision (IJCV)*, vol. 36, no. 2, pp. 149–161, 2000.

7. M.K. Chung, K.J. Worsley, J. Taylor, J.O. Ramsay, S. Robbins, and A.C. Evans, "Diffusion smoothing on the cortical surface," *NeuroImage*, vol. 13S, no. 95, 2001.

8. András Kelemen, Gábor Székely, and Guido Gerig, "Elastic model-based segmentation of 3d neuroradiological data sets," *IEEE Transactions on Medical Imaging (TMI)*, vol. 18, pp. 828–839, October 1999.

9. Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 23, no. 6, pp. 681–685, 2001.

10. L.H. Staib and J.S. Duncan, "Boundary finding with parametrically deformable contour models," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 14, no. 11, pp. 1061–1075, Nov 1992.

11. M. Leventon, O. Faugeraus, and W. Grimson, "Level set based segmentation with intensity and curvature priors," in *Workshop on Mathematical Methods in Biomedical Image Analysis Proceedings (MMBIA)*, June 2000, pp. 4–11.

12. L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever, "The Gaussian scale-space paradigm and the multiscale local jet," *International Journal of Computer Vision (IJCV)*, vol. 18, no. 1, pp. 61–75, April 1996.

13. G. Székely, A. Kelemen, Ch. Brechbühler, and G. Gerig, "Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models," *Medical Image Analysis*, vol. 1, no. 1, pp. 19–34, 1996.

14. A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, 2001.

15. SM Pizer, T Fletcher, A Thall, M Styner, G Gerig, and S Joshi, "Object models in multiscale intrinsic coordinates via m-reps," in *Generative-Model-Based Vision (GMBV)*, 2002.

16. Conglin Lu, Stephen M. Pizer, and Sarang Joshi, "A markov random field approach to multi-scale shape analysis," in *ScaleSpace*, 2003, vol. this volume.