Profile Scale-Space for Image Segmentation

Sean Ho *, Guido Gerig

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

Abstract. Object boundaries in images often exhibit a complex greylevel appearance, and segmentation which accurately fits the target image requires a robust, multiscale statistical model of image appearance around the object. Objects whose appearance differs from region to region call for an image appearance model which is tied to the geometric representation of the object, after the fashion of Active Appearance Models [1]. This paper aims at improving the image match component of a model-based segmentation framework. We extract 1D profiles normal to the boundary, a la Active Shape Models, and develop a scale-space on the profiles, where blurring is done only parallel to the boundary. A statistical model is built on features in the profile scale-space, incorporating weighted feature selection. This yields image forces which, when coupled with shape constraints, provide a Bayesian image segmentation framework.

1 Introduction

Segmentation tasks in medical imagery strongly benefit from model-based segmentation. First, anatomical objects are variable in shape, size and pose but can be represented by statistical shape models, obtained from a representative set of training data. Second, contrast and type of transition at boundaries can significantly vary at different locations of an object boundary but also at corresponding points across the object population, thus inhibiting the use of simple boundary models. Third, contrast to noise is often low and needs to be accounted for by strong model priors. However, unlike natural scenes, the domain of anatomical shapes is largely constrained and is therefore especially suited for modeling shape and appearance of objects.

The two components of Bayesian model-based segmentation are geometric typicality and local image match. The geometric typicality prior $p(m)$ is guided by a model of the expected geometry and its variability within a population. Similarly, the image match force $p(I|m)$ is guided by a model of the expected image appearance around the object. Much work has been done on generating geometric models; here we will focus on the image match model, assuming the

* Supported by NIH-NCI P01 CA47982.

presence of a suitable geometric model. This paper describes building a multi-scale model of local boundary properties and its use to generate local forces to drive segmentation. We put emphasis on a multi-scale framework to improve robustness, the possibility to deal with arbitrary profile shapes as common in medical images, and the potential for automatic selection of the appropriate scales that takes into account image noise, population variability, and object boundary complexity. We describe both, constructing of a function for $p(I|M)$ that can be sampled (Sections 4.3 and model fitting in segmentation by a framework to guide optimization (Section 4.4). The implementation used in this paper is in 2D, but Section 6 discusses extensions to 3D.

2 Related Work

There are various image match models in use by current model-based segmentation methods.

Gradient Magnitude Perhaps the simplest image match model is simply to optimize for large image gradient magnitude. This is often used by both classical mesh-based snakes [2] and implicit (geodesic) snakes [3]. The gradient magnitude model is global (assumes the boundary appearance is uniform across the whole object) and static (not statistically trained). Gradient magnitude image match performs poorly at weak or nonexistent step edges, or when the boundary discontinuity is not well-represented by a step edge.

Inside/Outside Classification Region-competition snakes use global probability distributions of “inside” intensities vs. “outside” intensities to drive the image match [4]. van Ginneken et al [5] perform a local inside/outside classification using not just the raw greylevels but a high-dimensional n-jet feature vector, with kNN clustering and feature selection, based on the training population. Leventon et al [6] train a global profile model that relates intensity values to signed distances from the boundary, incorporating more information than a simple inside/outside classifier.

Template Matching Segmentation via registration to a template image or atlas is another common technique. A single template serves as a standard, and segmentation performance depends on the similarity of the subject’s image appearance to the template.

Profile Model Cootes and Taylor’s seminal Active Shape Model work [7] samples the image along 1D profiles around boundary points, normal to the boundary, using correspondence given by the Point Distribution Model of geometry. At each point along the boundary, a probability distribution is trained on the image-derived features in the 1D profile (ASMs use the derivative of image intensity along the profile). The corresponding profile is sampled from the target image, and the Mahalanobis distance in the probability distribution at that boundary point provides the goodness-of-fit. 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. Profiles have also been used in 2D cardiac MR images [9].

Active Appearance Models Perhaps the most well-developed work on modeling intensity variation in objects is the Active Appearance Model [1], 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). To improve the optimization process, global PCA are performed on the registered images blurred at various scales. 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.

3 Open issues, Problems

Real world images and medical imagery can have different and complex image profiles at different locations along the boundary. For instance, an object might be expected to have a step edge at one part of the boundary, but a line-like edge at another part along the boundary. It is clear that gradient magnitude, especially at fixed global scale, is not sufficient to capture complex image characteristics. A robust image match model also needs to be trained from a population of images; a single template is not sufficient. Current profile models capture local variation, but do not take advantage of the expected high degree of correlation between adjacent profiles along the boundary. Normalization of the profiles is also an issue; selection of appropriate image-derived features is important. This work can be interpreted as an extension of Active Shape Models to capture local intensity variation in a multiscale fashion.

We would like a robust image match model able to capture local intensity-derived features about the boundary, while being aware of the distinction between the *along-boundary* and *across-boundary* directions, as defined by the geometric model. Along the boundary, we would like to take advantage of spatial locality in order to provide robustness to image noise and jitter in the provided correspondence. We would like an image match model which can capture image characteristics about the object boundary in a fashion robust to noise and efficient to train, in order to improve segmentation.

4 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 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.

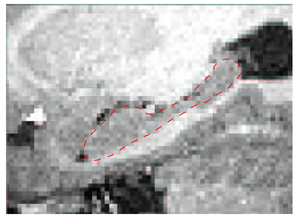


Fig. 1. Sagittal slice of the hippocampus, showing complex boundary appearance.

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.

4.1 Toy Example

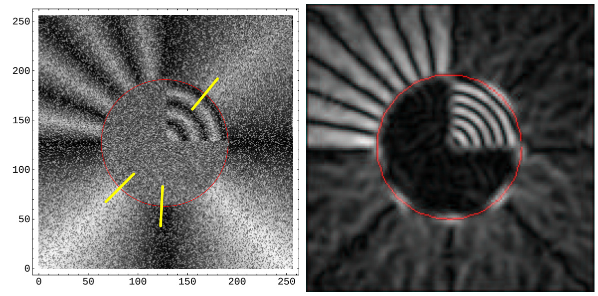
Throughout the discussion of our method, a toy example population of images will be used for ease of explanation, where the object shape is very simple (a circle in 2D), but the greyscale intensities are complex and have variability. We use 2D Fourier coefficients [10] as a parameterized shape representation; 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 from outside to inside the boundary; the profile at the bottom of the circle shows a weak 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.

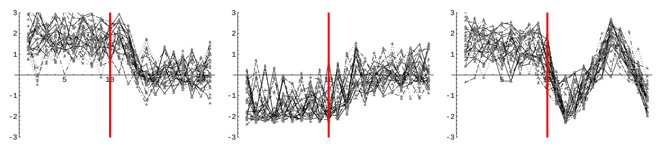
4.2 Scale-Space in Object-Intrinsic Coordinates

Our object shape representation (2D Fourier harmonics) provides a mapping from a standardized parameter space to the object boundary in each image in



(a) (b)

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.



(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.

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 [11].

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 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.

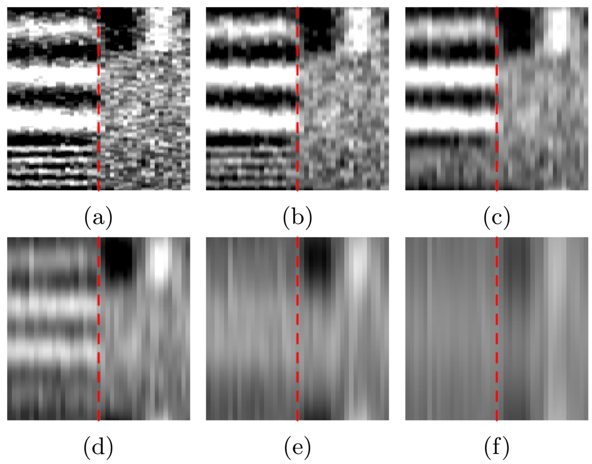


Fig. 4. The Gaussian scale-space constructed from one case in the toy example training population. Levels of the scale-space are (a)–(f), fine-scale to coarse-scale. Each row within a plot represents a boundary profile at that scale. The object boundary runs vertically through the center of each plot; left is exterior of the object and right is interior.

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 4.3. 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.

4.3 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:

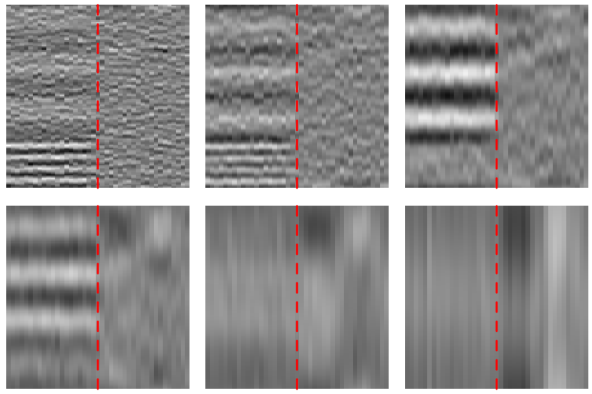


Fig. 5. The corresponding Laplacian scale-space. Levels of scale-space are shown fine-to-coarse, as in Figure 4. The coarsest level is duplicated from the Gaussian scale-space.

- 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.

With our toy example, each profile is 21 samples long, there are six scale levels, with 128 profiles around the boundary, and 30 images in the training population. The 21 samples along each profile are spaced in units of two pixels. At each combination of (σ, u) in the scale-space, we have a 21-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 top left 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 model is trained by many local low-dimensional PCAs linked in the Markov chain, instead of one global very high-dimensional PCA, as in e.g. Active Appearance Models. Thus 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 [6], 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.

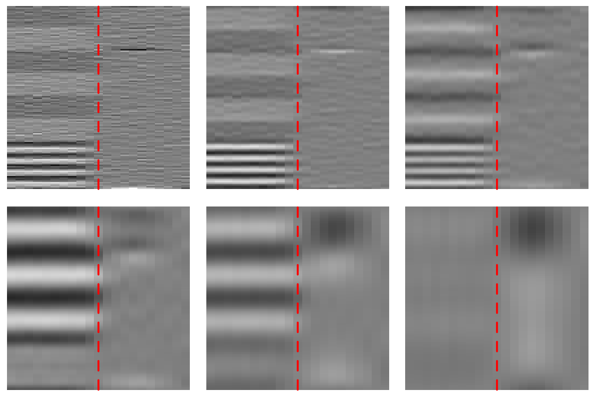


Fig. 6. Training population mean, Laplacian scale-space. Levels of the scale-space are shown fine-to-coarse, as in Figure 4.

4.4 Automatic Feature Selection

We would like to rate how useful a particular profile $I(\sigma, u)$ is for segmentation. If a profile has a complex but statistically stable edge structure, we want to weight it more heavily for use in segmentation, whereas if a profile has very large variability in the training population, we want to deweight it, even if the mean happens to show a nice step edge. Inspired by Canny's localization criterion for edge detection [13], we call this weighting *LOC*.

To use this image match model in a deformable model segmentation framework, the profiles would be extracted and scale-space built at each iteration of the segmentation, for the current stage of deformation of the geometric model. At each location and level of the scale-space, a goodness-of-fit measure can be evaluated from the Mahalanobis distance to the local Gaussian distribution. In addition, the goodness-of-fit can be evaluated for various shifts of the profile normal to the boundary, to drive the image forces to deform the geometric model to the optimal location.

At the model-building phase, the target image is unknown, but we can use the training population to rate the usefulness of profiles for segmentation. At each (σ, u), we evaluate the Mahalanobis distance of the mean profile, when shifted back and forth (left and right on the visualizations). Obviously, at zero shift, the Mahalanobis distance is always zero. However, the behavior for nonzero shifts depends on the local edge structure relative to the local variability (Figure 7). We would like a sharp dip to zero Mahalanobis distance at zero shift, with large Mahalanobis distances at any nonzero shifts. We fit a parabola to the plot of goodness-of-fit and define *LOC* (in the range [0,1]) in terms of the width of the parabola; there are certainly other ways to do it as well. The color coding in Figures 7 is by LOC; red to violet is higher weighting, and yellow to green to blue is lower (the color map is cyclic). Note also that LOC is not a binary feature selection but a continuous deweighting of irrelevant features. Of course, one could also threshold LOC to perform a binary feature selection.

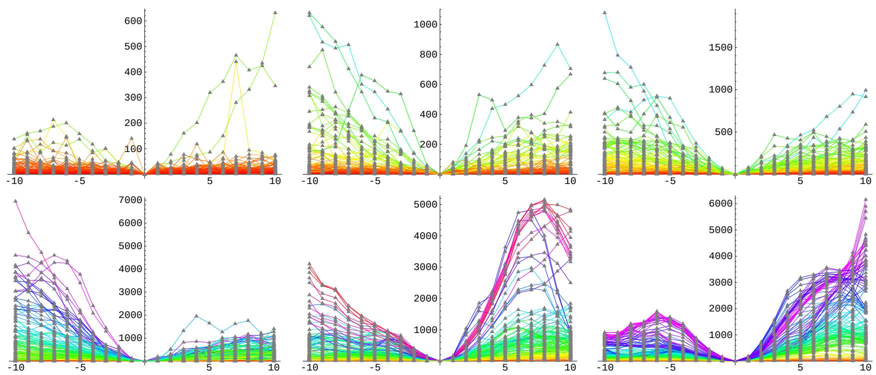


Fig. 7. Goodness-of-fit of the mean profile, when shifted against itself by ± 3 samples. Levels of the scale-space are shown coarse-to-fine, as in Figure 4. Each curve within a graph is a different sector of the boundary at that scale. Horizontal axis is shift of the profile normal to the boundary (zero shift is at center). Vertical axis is goodness-of-fit against the local statistical model. Color coding is a measure of sharpness of the goodness-of-fit graph (feature weighting): red to violet is a sharper graph more useful for segmentation; yellow to green to blue is a flatter graph.

5 Segmentation

At each iteration of segmentation, we extract the profiles normal to the currently deformed contour, and build the profile scale-space. At each profile at each scale in scale-space, the profile is compared to the model profile for various shifts of that profile normal to the contour. The shift which gives the smallest Mahalanobis distance is selected as the suggested shift for that profile at that scale. This yields a set of suggested shifts for every location around the snake. A weighted sum across scale using the LOC feature weighting is used to construct a single shift vector for each location around the snake. Finally, we reinterpolate to the Fourier shape representation (applying some smoothing in the process), apply shape constraints a la Active Shape Models, and iterate.

We compare our method to the single-scale profile model used in Active Shape Models. We use the same 2D Fourier shape model with both profile models, on synthetic data. The same initialization is used for both. Since the interaction between the shape model and the image match model in Bayesian segmentation can be complex, we show results without shape constraints (the interpolation to 2D Fourier harmonics still provides some smoothing). Use of a shape prior would aid the final segmentation, however in this paper our focus is on the performance of the image match model and not on the shape prior.

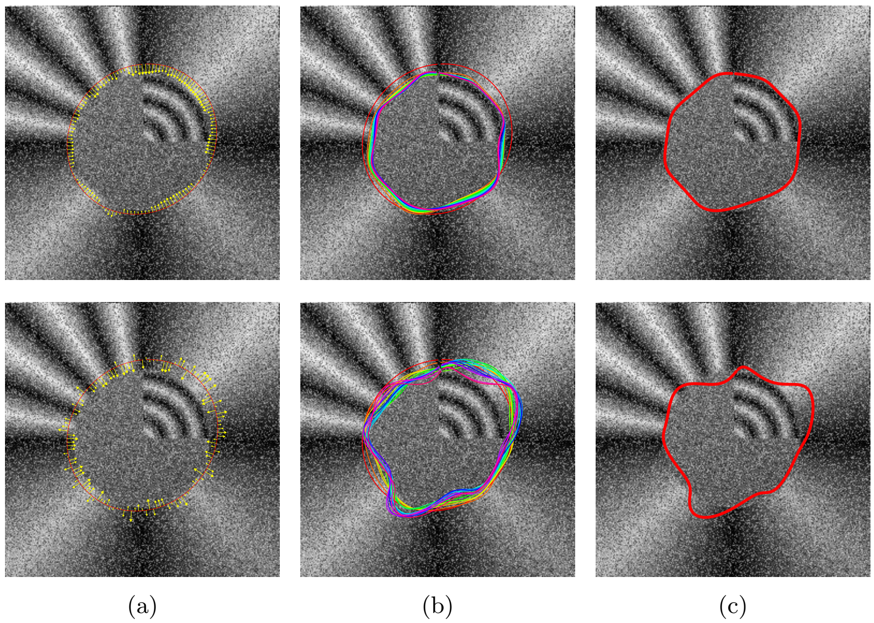


Fig. 8. Comparison of our scale-space profile model (top row) to single-scale profiles (bottom row), when used *without* any shape prior, highlighting the performance of the image match component of model-based segmentation. (a) Image forces at first step. (b) Overlay of 15 iterations of segmentation. (c) Result after 15 iterations.

The segmentation results shown in Figure 8 are a difficult test for both types of snake, since there is no prior shape model, the images have a high level of noise, and the boundary edge structure changes from place to place (correspondence is important). Despite these difficulties, the snake using our image match model (top row of Figure 8) clearly "snaps" to the right boundary at almost all locations around the circle. The bottom row shows the classical ASM-type snake with single-scale profiles; in the presence of such a high degree of noise, the image forces are noisy and weak, and the snake generally does not move in the right direction. These results highlight the strength of our profile scale-space as an image match model; when coupled with an appropriate shape prior, the segmentation results would be even better.

6 Extension to 3D

Work is in progress on extending the profile scale-space model building framework from 2D images to 3D, using spherical harmonics for the geometric representation [8]. The parameter space for the 2D Fourier harmonics is the unit circle; the parameter space for the 3D spherical harmonics is the surface of the unit ball. In 2D, profiles are positioned about the circle uniformly in arclength. In 3D, profiles can be positioned uniformly on the parameter space using an icosahedron subdivision of the sphere. In 2D, we blur along the boundary using a Gaussian kernel in arclength about the unit circle. To blur along the unit ball, the Laplace-Beltrami kernel can be used, similarly to Chung et al [14].

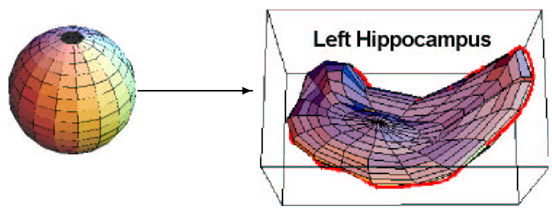


Fig. 9. Parameterized spherical harmonic representation of the hippocampus.

We also have an existing database of 85 MR images with expert hand segmentations of the left and right hippocampus and caudate, for training and testing of our new image match model. The hippocampus (Figures 10, 1) exhibits much more complex image characteristics about the boundary than the corpus callosum does. This motivates the exploration of multiscale image match models which leverage the geometrically-given correspondence.

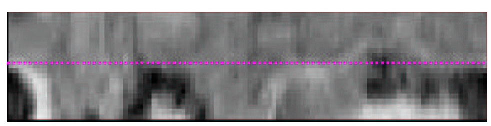


Fig. 10. Profiles taken about a meridian line of the hippocampus (shown in red on Figure 9). The hippocampus has a much more complex boundary profile structure than the corpus callosum, motivating an image match model that can capture complex local edge structure.

7 Conclusion

We present a novel statistical image match model for segmentation, using a scale-space in a coordinate system relative to the object boundary. Building on the classical notions of image pyramids, we apply the Laplacian scale-space to boundary profiles, thus blurring along the boundary given by geometry, while preserving across-boundary image structure. A statistical model is constructed from many independent Gaussian models, tied together in a Markov chain. Automatic feature selection across scale is included, and the whole model-building framework fits in the Bayesian deformable models segmentation paradigm by defining a probability distribution $p(I|m)$ for image match, which can be empirically evaluated to test how well a given deformed geometric model m fits a particular target image I.

We demonstrate results on test data using our image match model in a full segmentation framework; the results clearly indicate improvement over the single-scale ASM profile model. To effectively segment real-world objects in medical images, we need a powerful statistically-trained image match model that can capture complex image structure in a coordinate system relative to the object geometry. Our future research will embed this multi-scale boundary model into 2D and 3D model-based segmentation schemes and explore its performance.

References

1. 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.
2. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active shape models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, 1987.
3. H. Tek and B.B. Kimia, "Volumetric segmentation of medical images by three-dimensional bubbles," *Computer Vision and Image Understanding (CVIU)*, vol. 65, no. 2, pp. 246–258, 1997.
4. S. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation," in *International Conference on Computer Vision (ICCV)*, 1995, pp. 416–423.
5. Bram van Ginneken, Alejandro F. Frangi, Joes J. Staal, Bart M. ter Haar Romeny, and Max A. Viergever, "A non-linear gray-level appearance model improves active shape model segmentation," in *Proc. of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, 2001.
6. 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.
7. 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.
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. Nicolae Duta, Anil K. Jain, and Marie-Pierre Dubuisson-Jolly, "Learning-based object detection in cardiac MR images," in *International Conference on Computer Vision (ICCV)*, 1999, pp. 1210–1216.
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. 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.
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. 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.
14. 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.