

Hierarchical Statistical Modeling of Image Profiles About the Object Boundary

CVPR2003 submission paper ID: 194

Abstract— Object boundaries in images often exhibit a complex 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 require an image appearance model which is tied to the geometric representation of the object, after the fashion of Active Appearance Models [1]. By utilizing the correspondence given by a geometric representation in a region of the image around the object, a hierarchical coarse-to-fine framework can be constructed for image intensities which allows us to study their variability in a population and develop a statistical model of object appearance. This can be used to develop an image match model in segmentation and registration.

I. INTRODUCTION

Model-based image segmentation revolves around two forces: geometric typicality and image match. The geometric typicality prior $p(m)$ is guided by a model of the expected geometry of the object to be segmented. 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 presence of a suitable geometric model. The main idea of this paper is to motivate moving away from a global, acquisition-based coordinate system for viewing the image, and towards a local, geometry-intrinsic coordinate system for viewing the image relative to the given object geometry, in a region of interest about the object boundary. Given a common geometry-based coordinate system for a set of training images, a multiscale hierarchy of anisotropic filters oriented according to the local coordinate system provides a useful set of image-derived features from which one can construct a probabilistic model for $p(I|m)$.

The methods presented here focus on the model-building task, constructing a function for $p(I|m)$ that can be sampled. For model fitting in segmentation, there is a framework to guide optimization that is touched upon in Sections IV-D and VII. The implementation used in this paper is in 2D, but Section VI discusses extensions to 3D.

II. RELATED WORK

A. 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 disconti-

nity is not well-represented by a step edge.

B. 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 k NN clustering and feature selection, based on the training population. Their method works particularly well when the inside/outside distinction is represented by differing textures.

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. A model is built of a single representative profile, together with variability estimated non-parametrically with Parzen windowing.

C. Template Matching

Segmentation via registration to a template image or atlas is a common technique. For example, the medially-based M-reps deformable model segmentation framework [7] can use image intensities taken from a single template image to drive its image match forces. The correlation between corresponding intensities in the template image and in the target image provides the goodness-of-fit. This takes advantage of the correspondence defined by the geometric model. It is natural to wish to extend this to use a training population instead of a single template image.

D. Profile Model

Cootes and Taylor’s seminal Active Shape Model work [8] 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 [9] can be seen as a variant of ASMs, and uses a training population linked with correspondence from the geometric model, and 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.

Figure 1 shows an ASM-based automatic segmentation of the corpus callosum in 2D, using a 2D Fourier shape representation [10] [11]. There is a visualization of the greyscale profiles sampled around the boundary, as well as the goodness-of-fit of each profile to its local model as it is shifted by various increments normal to the boundary.

At initialization of the mean shape model into a new target image, the profile visualization clearly shows the mismatch into the image, as well as indicating suggested shifts to optimize towards a better match. At final segmentation, we see a clear step edge around the boundary. The bottom right image shows goodness-of-fit of the profiles, indicating a nice “valley” at zero shift (no further motion) across all the profiles.

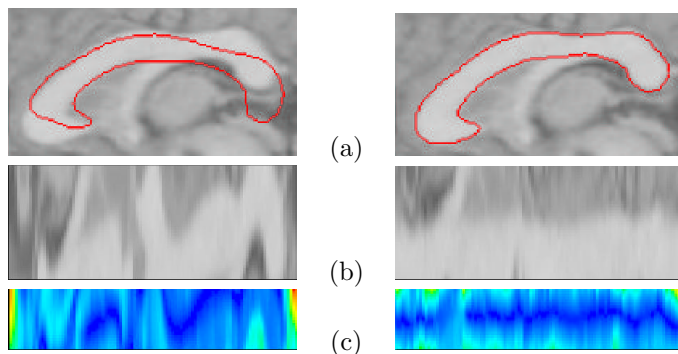


Fig. 1. Image profiles about the corpus callosum in 2D. Left: with a poor fit to the image. Right: with a good fit. (a) Fourier contour overlaid on image. (b) 1D greylevel profiles about the contour, unrolled. (c) Goodness-of-fit of each profile at various shifts normal to the boundary. Dark blue indicates better fit to the per-profile statistical model.

E. 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.

III. OPEN ISSUES, PROBLEMS

Real world images and medical imagery exhibit complex image profiles at the object boundary. Moreover, boundary profiles can have different characteristics 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. Active Appearance Models and Active Shape Models are very promising and closely related; this work can be interpreted as an extension of ASMs to capture local intensity variation in a hierarchical 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, already 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. There are many potential sources of image noise in the image acquisition process. For instance, in magnetic resonance images (MRIs), we face MR noise, scanner stability, and bias field inhomogeneity, in addition to patient induced variations of brightness and contrast, and inter-object variability (e.g. motion of adjacent structures). We would like an image match model which can capture image characteristics about the object boundary despite noise, in order to improve segmentation.



Fig. 2. Sagittal slice of the hippocampus, showing complex boundary appearance.

IV. METHOD

We propose viewing 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. Similarly to Canny’s [12] ideas and those of geometry-driven diffusion, we want to look for edges across the boundary, but smooth away noise along the boundary. To accomplish this, we construct an image pyramid, in a non-Euclidean object-intrinsic coordinate system which preserves the along-boundary versus across-boundary distinction.

A. 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 V 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 20 sample images were created.

Figure 3 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.

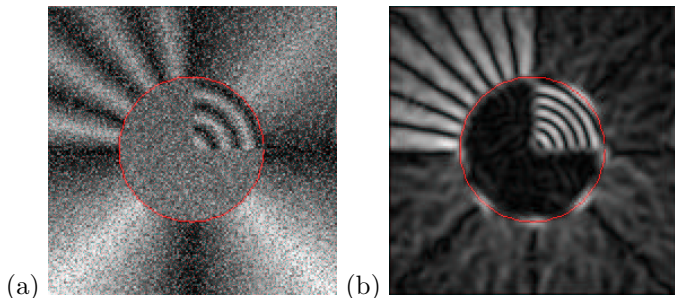


Fig. 3. (a) A sample from the toy example training population. (b) The corresponding gradient magnitude image at $\sigma = 4$ pixels, for reference. The contour from which profiles will be taken is shown in red. Each of the 20 images in the population have different additive noise, as well as multiplicative shading (linear ramp) in different directions.

B. Image Pyramid in Object-Intrinsic Coordinates

Our object shape representation 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 classical image pyramid in the original (x, y) image coordinates, edge structure across the boundary would also be blurred away. In the spirit of nonlinear diffusion and geometry-driven diffusion, we wish to blur along the boundary (to deal with noise), but detect edges across the boundary [13]. An elegant solution for diffusion along the object boundary is to map the blurring kernel from the parameter space onto the object.

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 y and x axes of the unrolled profile image, and we can construct a classical image pyramid in one direction, leaving the across-boundary direction unfiltered. The convolution in the

along-boundary direction is cyclic, since the parameter space is the unit circle. Thus, the top level of the Gaussian pyramid is a single average profile summarizing the whole boundary; the next finer level has two profiles each summarizing half the boundary, and so on down to the finest level, which is the original profiles sampled from the image. As a global brightness/contrast normalization pre-processing step, the profile images are scaled to have zero mean and unit variance, before the pyramids are constructed.

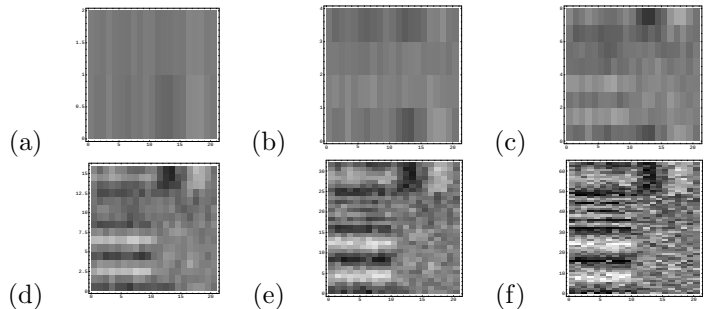


Fig. 4. A sample Gaussian pyramid from the toy example training population. Levels of the pyramid are (a)–(f), coarse-to-fine. Each row within a plot represents a summary profile about a sector of the boundary at that scale. The object boundary runs vertically through the center of each plot; left is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

We build the Laplacian pyramid ($\partial/\partial\sigma$) on the profile image, representing each level of the pyramid as a residual from the next coarser level. This sets up a Markov chain in scale-space, as will be discussed in Section IV-C. 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. We chose to implement the scale-space as a discrete pyramid; one could also build a continuous scale-space [14] representation on the profile images.

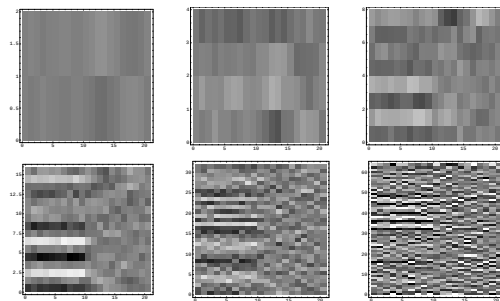


Fig. 5. The corresponding Laplacian pyramid. Levels of the pyramid are shown coarse-to-fine, as in Figure 4.

C. Local Statistical Model

The Laplacian pyramid 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 pyramid 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 pyramid,
- 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 15 samples long, there are six scale levels, with 64 profiles at the finest level, and 20 images in the training population. The 15 samples along each profile are spaced half a pixel apart. At each combination of (σ, u) in the pyramid (127 total), we have a 15-dimensional feature space and a cluster of 20 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 pyramid, 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 [6], could also be done instead of the Gaussian estimation.

By treating each profile in the pyramid 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 [14]. The training population “tells us” what kind of edge to look for locally, including variability.

Figure 6 shows the mean profiles at all locations and scale levels in the pyramid. 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 pyramid, 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 nice edge structure in the mean profile, however the variability due to noise at that scale overwhelms the step edge; compare with the pyramid of a single case in Figure 5.

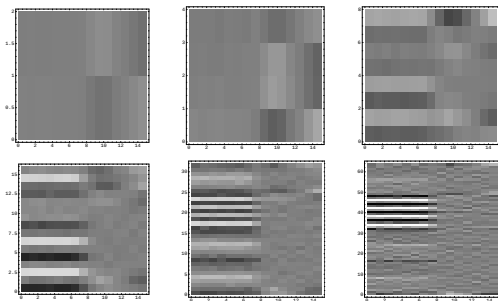


Fig. 6. Training population mean, Laplacian pyramid. Levels of the pyramid are shown coarse-to-fine, as in Figure 4.

D. 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 [12], we call this weighting *LOC*.

To use this image match model in a deformable model segmentation framework, the profiles would be extracted and pyramid built at each iteration of the segmentation, for the current stage of deformation of the geometric model. At each location and level of the pyramid, 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. Figure 8 shows goodness-of-fit in a leave-one-out test.

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 and 8 is by *LOC*; blue to purple to red is higher weighting, and green to yellow to red is lower (the color map is cyclic). Note also that *LOC* is a continuous rating; this 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.

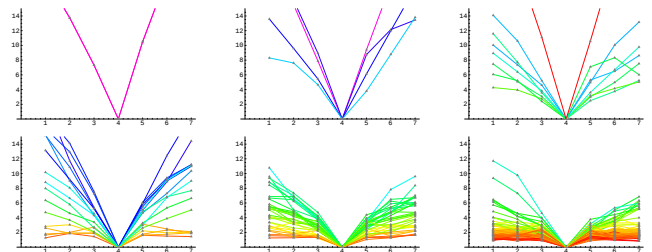


Fig. 7. Goodness-of-fit of the mean profile, when shifted against itself by ± 3 samples. Levels of the pyramid are shown coarse-to-fine, as in Figure 4. Each curve within a graph is a different sector of the boundary at that scale. Horizontal axis is shift of the profile normal to the boundary (zero 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): blue to purple is a sharper graph more useful for segmentation; green to yellow to red is a flatter graph.

To see how the *LOC* rating helps segmentation, we do a leave-one-out test with our toy example population. Profiles from one image are left out of the training population as a test case (profiles from the test image need to be a little longer to allow room to shift). At each level of the pyramid, profiles from the test case are then shifted against the trained model, and the goodness-of-fit is plotted, as shown in Figure 8.

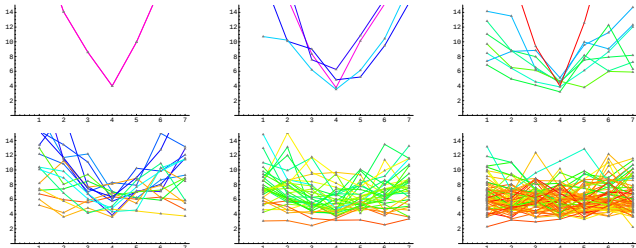


Fig. 8. Goodness-of-fit of a test profile to the statistical model in a leave-one-out experiment. Horizontal axis is shift in/out by ± 3 pixels. Axes are same as in Figure 7. Levels of the pyramid are shown coarse-to-fine, as in Figure 7.

In our toy example, the geometric model is always the same circle. Since the test profiles are taken relative to a perfectly fit geometric model, we expect the optimum Mahalanobis distance to be at zero shift. However, due to the large amount of image noise, the plots of goodness-of-fit are very noisy at fine scales, and the optimum may deviate significantly from the ideal zero shift. In those cases, the low value of *LOC* deweights those shifts, giving more weight only to the stable profiles at coarser scale. Figure 9 shows a scatterplot of suggested shift versus *LOC*. The ideal zero shift is the center line. We see that for many (σ, u) the suggested shifts are not optimal, but *LOC* is also low. The profiles where *LOC* is above 0.6 all suggest shifts of less than half a sample (a quarter pixel) from the ideal.

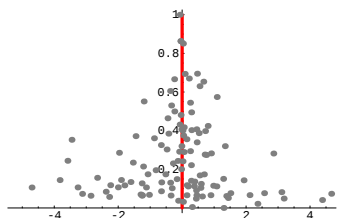


Fig. 9. Scatterplot showing interaction between feature weighting (*LOC*) and suggested shift in the leave-one-out experiment. Each profile at each scale in the pyramid is represented by a point; horizontal axis is the shift suggested by the local model (centerline is ideal), vertical axis is feature confidence.

V. APPLICATION: CORPUS CALLOSUM

We have built a statistical profile pyramid model on 71 segmented 2D corpus callosum images, for use in a segmentation framework with a shape representation of 2D Fourier harmonics. Each of the 71 corpora callosa have been hand-segmented by experts, and the Fourier harmonics constructed. Profiles at the finest scale of the pyramid

were extracted from the images (100 profiles at finest level), and the Laplacian pyramid built for each image. The large scale edge structure of the corpus callosum can indeed be represented by a single step edge, which is in the very top level of the Gaussian pyramid, as shown in Figure 10.

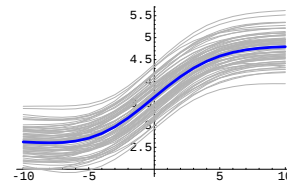


Fig. 10. Top level of the corpus callosum Gaussian profile pyramid (i.e. summary profile about the whole boundary). Vertical axis is greylevel intensity, horizontal axis is distance normal to the boundary (outside to the left; inside to the right). Population mean is shown overlaid on graph showing population variability.

Figure 11 shows the mean profiles of the Laplacian pyramids. As before there is also covariance structure behind each profile at each scale level that is not shown. The lower levels of the Laplacian pyramid 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 5 and 6 of the Laplacian profile pyramid, 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 1(a). 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) (Figure 10).

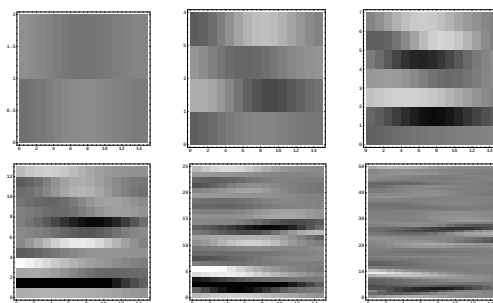


Fig. 11. Population mean of profiles around the corpus callosum; Laplacian pyramid. Levels of the pyramid are shown coarse-to-fine, as in Figure 4.

VI. EXTENSION TO 3D

Work is in progress on extending the profile pyramid model building framework from 2D images to 3D, using spherical harmonics for the geometric representation [9]. 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, a wrapped Cauchy basis function or vonMises basis function is a reasonable blurring kernel. [15]

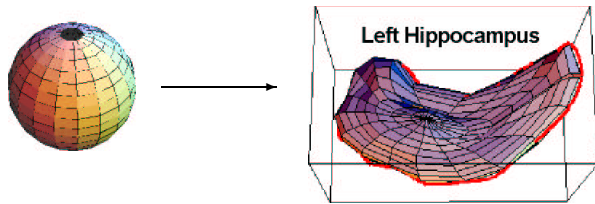


Fig. 12. Parameterized spherical harmonic representation of the hippocampus.

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 13, 2) 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.

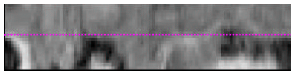


Fig. 13. Profiles taken about a meridian line of the hippocampus (shown in red on Figure 12). The hippocampus has a much more complex boundary profile structure than the corpus callosum, motivating an image match model that can capture complex local edge structure.

VII. CONCLUSION

We present a novel statistical image match model for segmentation and for image understanding in a coordinate system relative to the geometric object boundary. Building on the classical notions of image pyramids, we apply the Laplacian pyramid to boundary profiles, thus blurring along the boundary externally 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 on a continuous spectrum 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 .

In addition, the Markov structure of the joint statistical model provides a framework for a smart guided optimization of $p(I|m)$ in segmentation. 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 in order to optimize in the global Gibbs distribution, one can optimize in each of the smaller local Gaussian distributions. Each profile at each scale level of the pyramid can suggest a local deformation of the geometric model, together with a confidence rating in *LOC*. 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*, 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'95)*, 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*, 2001.
- [6] M. Leventon, O. Faugeras, 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] S. Pizer, D. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 851–865, October 1999.
- [8] 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.
- [9] András Kelemen, Gábor Székely, and Guido Gerig, "Elastic model-based segmentation of 3d neuroradiological data sets," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 828–839, October 1999.
- [10] L.H. Staib and J.S. Duncan, "Boundary finding with parametrically deformable contour models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1061–1075, Nov 1992.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] K.V. Mardia, *Statistics of Directional Data*, Academic Press, 1972.