

Shape Priors for Segmentation of the Cervix Region within Uterine Cervix Images

Shelly Lotenberg , Shiri Gordon and Hayit Greenspan.

Tel Aviv University, Tel-Aviv 69978, Israel

ABSTRACT

The work focuses on a unique medical repository of digital Uterine Cervix images (“Cervigrams”) collected by the National Cancer Institute (NCI), National Institute of Health, in longitudinal multi-year studies. NCI together with the National Library of Medicine is developing a unique web-based database of the digitized cervix images to study the evolution of lesions related to cervical cancer. Tools are needed for the automated analysis of the cervigram content to support the cancer research. In recent works, a multi-stage automated system for segmenting and labeling regions of medical and anatomical interest within the cervigrams was developed. The current paper concentrates on incorporating prior-shape information in the cervix region segmentation task. In accordance with the fact that human experts mark the cervix region as circular or elliptical, two shape models (and corresponding methods) are suggested. The shape models are embedded within an active contour framework that relies on image features. Experiments indicate that incorporation of the prior shape information augments previous results.

Keywords: Shape-Prior, Level Set function, Active Contour, Uterine cervix, Cervical cancer, Image segmentation and indexing

1. INTRODUCTION

The purpose of this work is to embed prior shape information into the task of cervix region segmentation within uterine cervix images (“cervigrams”). The images in this work were collected by the National Cancer Institute (NCI) in order to study the evolution of cervical cancer.¹ The work is part of a multi-stage system that is being developed in order to automatically analyze the content of the cervigrams within this database.²

Some typical cervigrams are presented in Figure 1. The cervix region (marked by a blue line), is located in the central part of the image, with surrounding vaginal walls and parts of clinical equipment, intruding in the image. A dark surrounding frame can be seen containing lines and text that are overlaid on the image at time of the photographic development. The cervix region is defined by the cervix boundary. Automated detection of the cervix boundary is an important preprocessing step, which defines the region of medical and anatomical interest within the cervigram and enables further analysis to focus within the cervix region itself.

In previous works, the detection of the cervix region was performed based on features derived from the input image. An initial rough region of interest (ROI) was initially extracted using features of color and position.² A more accurate delineation of the boundary was then enabled using an active contour framework based on curvature and color features.³ The current work extends the active contour framework by integrating prior shape information into it. The anatomical region features of the cervix (as marked by the medical experts) can be characterized by their elliptical or circular shapes, hence the ellipse and the circle are chosen for the shape models. Embedding shape information into the cervix region segmentation process is shown in the current work to improve the segmentation quality.

Address all correspondence to H. Greenspan, E-mail:hayit@eng.tau.ac.il

Copyright 2008 Society of Photo-Optical Instrumentation Engineers. This paper will be published in SPIE Medical Imaging Symposium and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple production, distribution to multiple locations via electronic or other means, duplications of any material in this paper for fee or commercial purpose, or modification of the content of the paper are prohibited

A color version is available at http://www.eng.tau.ac.il/shiri/mip_lab/

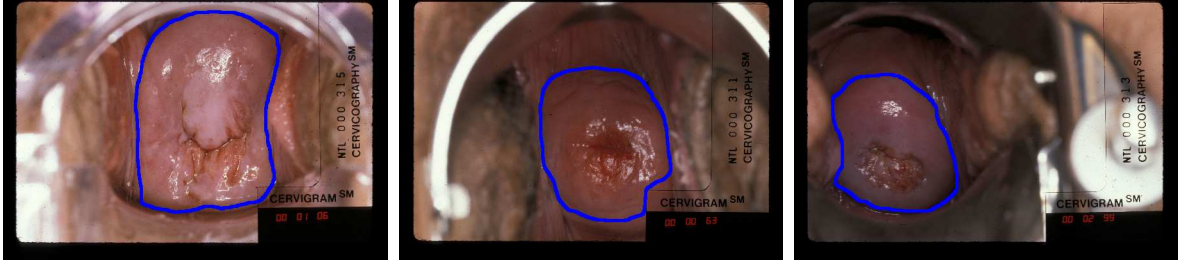


Figure 1. Cervix region examples marked by a medical expert.

A vast amount of work had been done to embed prior-shape information into a segmentation task. A popular approach is to use prior models based on allowable deformation of a template shape. In this approach the shape model is generated from an aligned training-set and different tools are used for its representation. Cootes *et al.*⁴ presented the Active Shape Model, which uses the statistics of a large set of points, sampled at meaningful locations along the object boundary. The generated high-dimensional model requires a vast amount of computational power, thus dimensionality reduction is desired. Leventon *et al.* used Principle Component Analysis (PCA) in order to extract the mean shape and the eigenshapes and represented them with a Gaussian distribution.⁵ In a different work of that group, the model is represented by a level-set function, using intensity and curvature as a function of the signed distance from the object boundary.⁶ Following model generation there are works that use an active contour framework in a level-set implementation for the segmentation task.^{7,8} In these works, registration and segmentation are performed simultaneously to align a new image to the prior model. Methods, such as those mentioned above, are inappropriate for the cervix region segmentation task due to the large variability of cervix shapes that exists across the database. Such variability leads to a model with a large range of allowable deformations, that has a minor effect on the resulted segmentation.

The objective of the current work is to propose a method that can cope with the large variability of shapes and can incorporate elliptical or circular shapes information in a straight-forward and a low-complexity scheme. Two methods are presented:

- A method that utilizes a histogram of the curve features with a circle shape model (presented in section 2.1).
- A novel method that uses a distance measure between level-sets and an ellipsoid-based shape model (presented in section 2.2).

2. METHOD

In an active contour framework the image is considered as a function $I : \Omega \rightarrow R^+$ where $\Omega \in R^2$ is the image domain. The segmentation problem is mathematically formulated as searching for a contour $C : [0, L] \rightarrow R^2$ in the image, which is optimal with respect to some pre-defined integral measure, $E(C)$, also called the energy functional. Formally this problem is stated as: $C = \operatorname{argmin}_c E(C)$. In the current work the energy functional consists of two terms: a data term and a shape prior term:

$$E(C) = E_{data}(C) + \alpha E_{shape}(C) \quad (1)$$

The data term, developed in a former work,³ is activated first and evolves the curve according to the image region and curvature-based edge properties. The prior-shape term is added next and better aligns the contour to a predefined model of the cervix's shape. The α parameter is a time dependant parameter that controls the activation sequence of the two terms.

2.1. Embedding a Circular Prior Using the Distribution of Shape Features

The first method for embedding shape information in the active contour framework is based on the distribution of shape features.⁹ The shape based energy term within the active contour functional penalizes the difference between a set of feature distributions of a given curve and those of a prior reference curve. It can be shown that such distributions can capture the intuitive similarity of shapes in a flexible way, while being invariant to shape transformations.¹⁰ This method is applied here with a circular shape prior and is hereon termed the *circular prior*.

We start with a brief description of the method. We then present specific modifications used for the cervix region segmentation task. The shape distribution is defined as the cumulative distribution function (CDF) of feature values measured uniformly along the shape boundary. Let Φ be a continuously defined feature along the curve C , and λ be a variable spanning the range of values Λ of the feature, the CDF of Φ , $H(\lambda)$, is defined as:

$$H(\lambda) = \frac{\int_C h\{\Phi(C) < \lambda\} dw}{\int_C dw}. \quad (2)$$

Here $h(x)$ is an indicator function, which is 1 when the inequality is satisfied and 0 otherwise. When it is meaningful to exhibit the particular curve C for which $H(\lambda)$ is computed, we will write $H(C, \lambda)$.

The shape-based energy term, $E_{prior}(C)$, is defined as:

$$E_{prior}(C) = \int_{\Lambda} [H^*(\lambda) - H(C, \lambda)]^2 d\lambda, \quad (3)$$

where $H(C, \lambda)$ is the feature distribution function of the curve C , and the prior shape information is captured in the target distribution $H^*(\lambda)$.

The shape descriptor used in the current work to describe the circular shape prior is termed the “inter-node distances” descriptor. This descriptor captures the CDF of the normalized distances between all nodes within the set of nodes S , sampled uniformly along the curve. The curve evolution minimization equation for the inter-node distances feature is⁹:

$$\nabla E(\Gamma)(s) = 2 \int_{t \in S} n(s) \cdot \frac{\vec{\Gamma}(s, t)}{|\vec{\Gamma}(s, t)|} [H^*(F_{st}) - H(\Gamma, F_{st})] dt, \quad (4)$$

where Γ is the parameterized curve as a function of the arc-length $\{X(s), Y(s)\}$ with $s \in \{0, 1\}$, $\vec{\Gamma}(s, t)$ is a vector with coordinates $\{X(t) - X(s), Y(t) - Y(s)\}$ and $n(s)$ is the outwards normal at $\{X(s), Y(s)\}$. The normalized inter-node distance between nodes (s, t) , is defined as:

$$F_{st} = \frac{|\vec{\Gamma}(s, t)|}{\text{mean}(\{|\vec{\Gamma}(s, t)| \mid (s, t) \in S\})}. \quad (5)$$

In the current work, the level set formulation for the shape term is implemented as:

$$\phi_t^{shape} = \nabla E(\Gamma)(s) * G_{\sigma}(x, y), \quad (6)$$

where $\nabla E(\Gamma)(s)$ is computed for pixels along the zero level set (the evolving curve C) and is diffused using a simple gaussian filter (G_{σ} in Equation (6)) of size 5×5 and $\sigma = 0.5$.

This method has various advantages that serve the current application: There is no need for registration and alignment of the prior-shape model and the evolving curve as the method is based on the distance between normalized feature distributions. These distributions are invariant to scale and rotation. On the other hand, the method is based on a non-parametric sampled curve, and its level-set implementation is not straightforward. Considering an image of size $M * N$ pixels, and assuming a sampled curve of size $O(x)$ where $x < M * N$, the algorithm complexity is $O(x^2)$. Originally this method was designed to deal with much more complex prior-shapes, and as such when applied to the cervigram region segmentation with a prior shape of a simple circle, it might seem too complex.

2.1.1. Combination of the data and shape terms in the curve evolution process

The data and shape terms, are combined in a two stage procedure: The first stage utilizes the data term to evolve the initial rough ROI to better fit the cervix region. The curve that is output from this stage will be termed as the *data-driven curve*. The second stage combines the data and shape terms to further refine the detected boundary and obtain a smoother, more circular shape.

This two-stage procedure was found necessary by experimentation: if the two terms are used simultaneously to evolve the curve in a single-stage procedure, the resulting cervix boundary is larger than desired: The data term is attracted to local concavities which are present between the initial ROI and the actual cervix boundary. The shape term tries to generate the optimal bounding circle that includes these concavities, thus preventing the curve from following the data inwards in favor of a more circular shape. A comparison between the two-stage procedure and the simultaneous activation of the data and shape terms, in a single stage, is presented in Figure 2. Figure 2(a) shows the medical expert-marked boundary (blue). Figure 2(b) shows the initial ROI (green), along with the boundary detected using data and shape terms in a single-stage procedure (red) and the boundary detected using the two-stage procedure (white). Local concavities can be found between the initial ROI curve and the desired cervix boundary (bright values in Figure 2(c)). The curve generated by the single stage procedure (red) is attracted to the concavities that are closest to the initial ROI, with the shape term enforcing a circular shape based on these concavities. It was empirically found that parameter tuning of within-term forces and between the two terms does not improve this outcome. The suggested two-stage procedure (white curve) results in a curve that matches more closely with the expert markings.

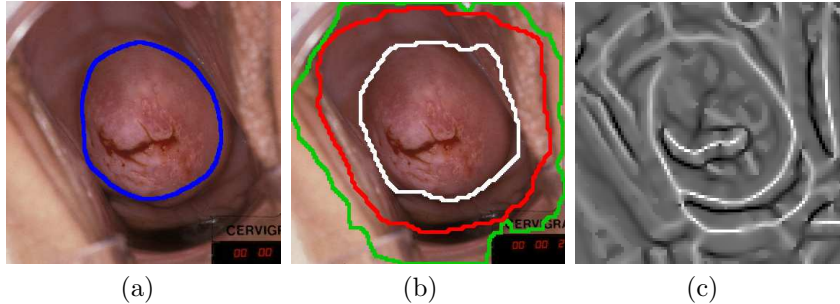


Figure 2. Incorporating prior shape information in the curve evolution functional. (a) manual marking of the expert (blue); (b) cervix boundary results: initial ROI (green); boundary detected using data and shape terms in a single-stage procedure (red); boundary detected using the two-stage procedure (white); (c) curvature feature map.

An important issue is the relative weighting of the data and shape terms in the second stage of the curve evolution process. Experiments have shown that equally weighting the two terms causes the data-driven contour to overinflate as it tries to match the circular prior model. In order to avoid this undesirable effect, the contour is restricted from evolving outwards beyond a pre-defined distance limit from the data-driven curve (which is in good proximity to the desired cervix boundary). This is done by using the following equation:

$$\phi_t = \phi_t^{data} + w_d(x, y)\phi_t^{shape}, \quad (7)$$

where ϕ_t^{data} is the level set formulation of the data term,³ ϕ_t^{shape} is defined by Equation (6) and the parameter, $w_d(x, y)$, locally weights the shape term per pixel. The local weights, $w_d(x, y)$, are defined using the signed distance transform, $d(x, y)$, that computes the minimal distance between a pixel (x, y) and the data-driven curve (positive distances interior to the curve and negative distances exterior to the curve). The weighting function, $w_d(x, y)$, is defined as:

$$w_d(x, y) = \begin{cases} \frac{N}{|d(x, y)|} & \text{if } d(x, y) < T \\ N & \text{otherwise,} \end{cases} \quad (8)$$

where T specifies the distance limit and is assigned a small value of $T = -3$ pixels, and N is a parameter that controls the general influence of the shape term ($N = 4$ in this work). The proposed weighting function

suppresses the influence of the shape term on pixels positioned outside the data-driven curve. Note that the weights are computed once per data-driven curve.

Figure 3 illustrates the effect of the local weights on the boundary detection quality. The data-driven curve is marked in green. The result for equally weighting the shape term is marked in red. The result for locally weighting the shape term, is marked in white. The manual markings of the expert are presented in blue. It can be observed that the equally-weighted curve (red) inflates the input data-driven curve (green). The locally-weighted curve (white) is smoother and better resembles the expert markings. This curve is shown to outperform the data-driven curve, thus supporting the inclusion of the shape term in the curve evolution process.

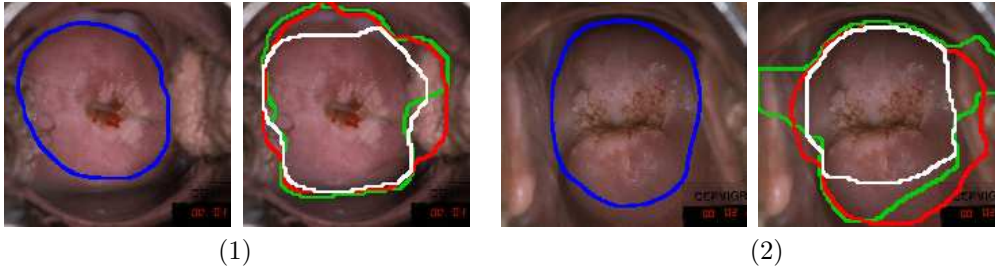


Figure 3. Local weights effect on boundary detection quality, two cervigrams examples. Left image - manual markings of the expert (blue). Right image - boundary detection results: Data-driven curve (green); Equally weighted shape term (red); Locally weighted shape term (white).

2.2. Embedding an Ellipsoid-Based Prior Using a Distance Measure Between Level-Sets

In this section, a second, novel method for embedding a shape prior into the cervix region segmentation task is presented. This method consists of two successive steps, computed in each iteration of the curve evolution process. The first step estimates an ellipsoid-based prior shape model, while the second step advances the evolving curve using a similarity measure between the level-sets of the curve and the model. This process simultaneously estimates the cervix boundary and the most adequate prior model for its representation. This method is hereon termed the *ellipsoid-based prior*.

The ellipsoid-based prior is generated using the bounding and bounded ellipsoids of the curve. It is computed using the following procedure: 1) The bounding ellipsoid is computed using the Geometric Bounding Toolbox of Matlab. 2) The bounded ellipsoid is computed using the covariance matrix of the coordinates within the contour area. An initial bounded ellipsoid is generated using a 2D projection of the covariance matrix at a fixed sigma. This initial projection is enlarged until an intersection with the curve is reached. 3) A set of shortest paths (lines) between the pixels of the bounding and bounded contours, is extracted. 4) The ellipsoid-based prior is defined as the centerline crossing the lines within this set.

The shape term of the active contour framework is derived from an integral of the non-overlapping areas between the zero level-set of the ellipsoid-based prior and the evolving curve. The integral measure is defined as¹¹:

$$E_{shape}(\phi) = \int_{\Omega} (H_{\epsilon}(\phi(x)) - H_{\epsilon}(\tilde{\phi}))^2 dx, \quad (9)$$

where $\tilde{\phi}$ is the prior shape level-set. H is the Heaviside function of a given level-set and it is used as an indicator function for the object and background regions in the image. H_{ϵ} is the a smoothed approximation of the Heaviside function¹²:

$$H_{\epsilon}(\phi) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan\left(\frac{\phi}{\epsilon}\right) \right). \quad (10)$$

The curve evolution equation is obtained by minimizing E with respect to Φ :

$$\phi_t^{shape} = \delta_{\epsilon}(\phi)(H_{\epsilon}(\phi) - H_{\epsilon}(\tilde{\phi})), \quad (11)$$

where δ_ϵ is approximated by:

$$\delta_\epsilon(\phi) = \frac{dH_\epsilon}{d\phi} = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + \phi^2}. \quad (12)$$

The shape term is weighted using $w_d(x, y)$ as in the case of the circular prior (Equation (7)).

This method is much simpler to implement (as compared to the first method) due to its level-set formulation. For an image of size $M * N$ pixels and curve of size $O(x)$, the algorithm complexity is $O(x)$. This complexity is attained by using a narrow band implementation of the level-set function. The model estimation phase of the current method is simpler as compared to other methods that estimate general-shape priors.¹¹ The ellipse based shape prior can be computed directly from the curve coordinates, as described above, thus avoiding the complex search for the optimal model transformation parameters.

3. EXPERIMENTS & RESULTS

The curve evolution framework presented in this work was evaluated using two sets of cervigrams for which the manual markings of a medical expert are available (120 cervigrams in *Set₁* and 158 in *Set₂*). The expert has a specialized medical experience in gynecological oncology. The cervigrams were randomly selected out of the NCI database (containing 100,000 images), without any restricting rules.

Experiments were conducted following preprocessing in which initial ROI was extracted and specularities were removed.^{2,13} Quantitative evaluation was carried out by comparing the generated contours to the expert-based markings. Several overlap metrics were used, including the Dice metric: $2 \frac{S \cap R}{|S| + |R|}$, the Sensitivity metric: $\frac{S \cap R}{R}$ and the False Positives (FP) metric: $\frac{S \cap \hat{R}}{\hat{R}}$, with S being the area of the automatically segmented region, R the expert segmentation and \hat{R} its complement (the area outside the expert marked boundary). The Hausdorff distance, a maximum surface distance,¹⁴ was used as well. The Hausdorff distance defines the largest difference between two contours (measured in pixels) and is a good indicator for shape resemblance, with a smaller Hausdorff distance corresponding to more similar contours.

3.1. Circular Prior Performance Evaluation

In the first experiment the influence of the circular prior on the segmentation performance is examined and the refinement of the initial data-driven curve is evaluated. Tuning the different parameters within the active contour framework was performed on a small training set of five cervigrams that were randomly selected out of *Set₁*. Once defined, the parameters were kept constant across all images in the test sets.

Figure 4 displays examples of cervix boundary detection. The data-driven contour generated in the first step of the curve evolution process is marked in green and the final result, following the refinement with the circular prior, is marked in red. Manual markings of the expert are imposed in blue. Corresponding overlap and Hausdorff measures are listed under each example. It is important to note when examining the Hausdorff results that the images are of approximately $1,500 \times 2,500$ pixels.

The following observations can be made from the results in Figure 4: The contours are smoother and more convex when the shape term is added, as desired. They exclude more irrelevant regions within the cervix. Good similarity to the expert markings can be seen in most of the cases. These facts are reflected by the quantitative evaluation, where a significant improvement in the FP and the Hausdorff metrics is achieved. The sensitivity results of examples (a)-(e) attain high values, as they include most of the expert's markings. These sensitivity results are similar for the two contours being compared. Example (f) shows an extreme case, where the sensitivity decreases due to the shape prior constraint. In this example the data-driven curve is already close, or located within the markings of the expert. Advancing the curve further using the circular shape prior generates a final smaller contour, as there are no color or curvature features that can prevent the curve from shrinking.

Table 1 presents the average cervix boundary segmentation results with (Circular) and without (No-prior) the shape priors, computed for the images of *Set₁* and *Set₂*. The Dice metric reflects a small improvement when using the circular prior. Sensitivity values are similar. These results indicate that the data-driven curve is already in good proximity to the desired contour and that regions within the cervix are handled well. The Hausdorff metric reflects a substantial improvement when applying the circular prior, thus indicating on a better

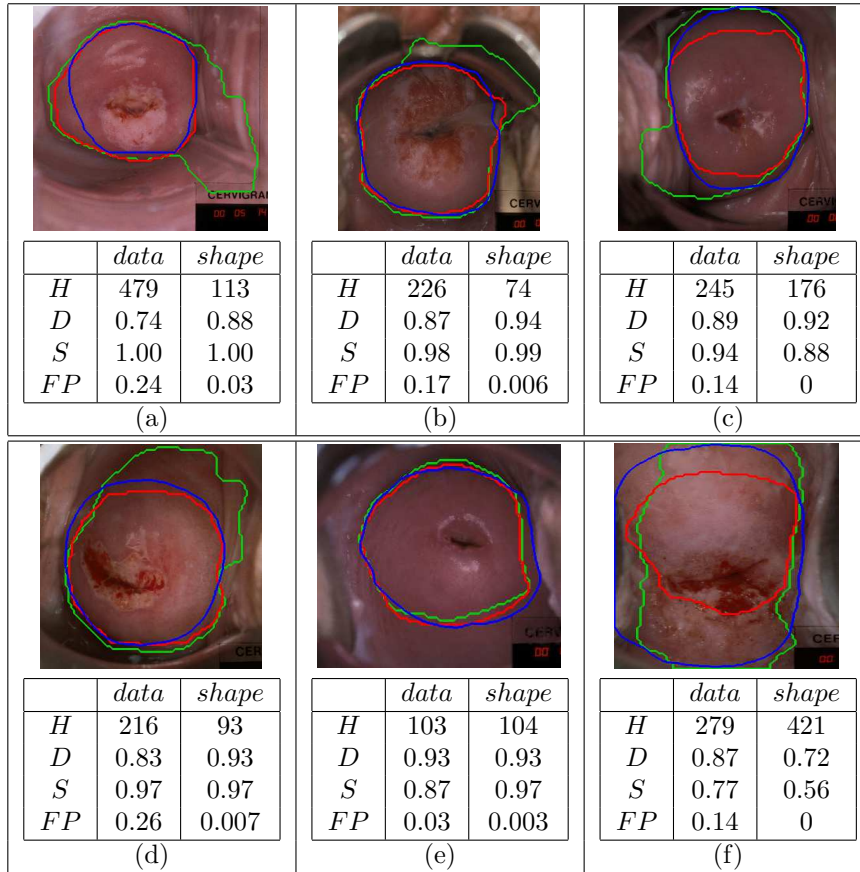


Figure 4. Cervix boundary detection. Data-driven contour, based on color and curvature features - marked in green; Final contour, following refinement with a circular prior - marked in red; Expert markings imposed in blue. Hausdorff (H), Dice (D), Sensitivity (S) and False Positives (FP), for the data-driven (*data*) and the circular shape prior (*shape*) contours, are listed under corresponding cervigrams.

shape resemblance of the generated contours to the expert’s markings. The FP metric reflects a considerable improvement, corresponding to a strong reduction of non-relevant tissues (outside the cervix region). This reinforces the addition of the circular prior to the curve evolution process. Results analysis is similar for both sets.

	<i>Set₁</i>			<i>Set₂</i>		
	No prior	Circular	Ellipsoid-based	No prior	Circular	Ellipsoid-based
Dice	0.79(0.10)	0.81(0.10)	0.82(0.09)	0.73(0.12)	0.75(0.13)	0.76(0.12)
Sensitivity	0.95(0.05)	0.94(0.08)	0.92(0.09)	0.96(0.07)	0.95(0.09)	0.93(0.1)
FP	0.28(0.13)	0.13(0.12)	0.11(0.1)	0.33(0.15)	0.16(0.14)	0.13(0.11)
Hausdorff	254(105)	216(91)	210(87)	300(125)	231(115)	224(101)

Table 1. Average results of cervix boundary segmentation with and without the shape priors. Mean value and standard deviation (within parentheses) are presented over the two data-sets used for evaluation (*Set₁*, *Set₂*).

3.2. Ellipsoid-Based Prior Performance Evaluation

The second experiment examines the performance of the Ellipsoid-based prior. Tuning the different parameters within the active contour framework was performed on a small training set of five cervigrams, as in the former experiment, and the parameters were kept constant across all images in the test sets.

Figure 5 shows example results of the two shape priors. Manual markings are imposed in blue, results of the circular prior are imposed in red and results of the ellipsoid-based prior are imposed in white. Examples (a) and (b) show cases where cervix boundary detection by the circular prior is more accurate. In these cases there is a high visual resemblance of the manual markings to a circle and the circular model is more appropriate. The cervix boundary in examples (c) and (d) has a high resemblance to an ellipsoid. The ellipsoid-based prior achieved better results in these cases. Average results of the ellipsoid-based prior, computed over the images of *Set₁* and *Set₂*, are presented in Table 1 (Ellipsoid-based). A comparison between the results of the ellipsoid-based prior and the circular prior reveals similar Dice and Sensitivity values. The FP and the Hausdorff metrics show a slight improvement when using the ellipsoid-based prior as compared to the circular prior. Overall the results computed over the entire data-sets are similar.

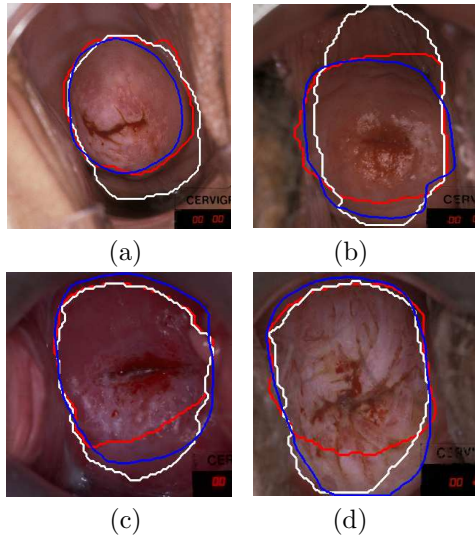


Figure 5. Cervix boundary segmentation with different shape priors. Manual markings imposed in blue; circular prior results imposed in red and ellipsoid-prior prior results, imposed in white.

In the final experiment we pursue a more thorough investigation of the two shape priors. We try to isolate the influence of the shape-prior term on the evolving curve and reduce the effect of other factors that might interfere with the curve evolution process. Two special sub-sets of cervigrams were generated for this experiment, each containing 31 images. The first subset includes cervigrams in which the manual expert markings resemble a circle (the circular sub-set). The second sub-set includes cervigrams in which the markings resemble an ellipsoid (the elliptical sub-set). The data-driven curve obtained in the selected images, is already in good proximity to the manual markings of the expert, thus a good initialization for the second step of the curve evolution process is enabled. This good initialization prevents the curve from converging into local optimum, which are related to the image features and are located outside the actual cervix boundaries. The curve evolution process, in the selected cases, is mainly driven by the shape-prior term, as desired.

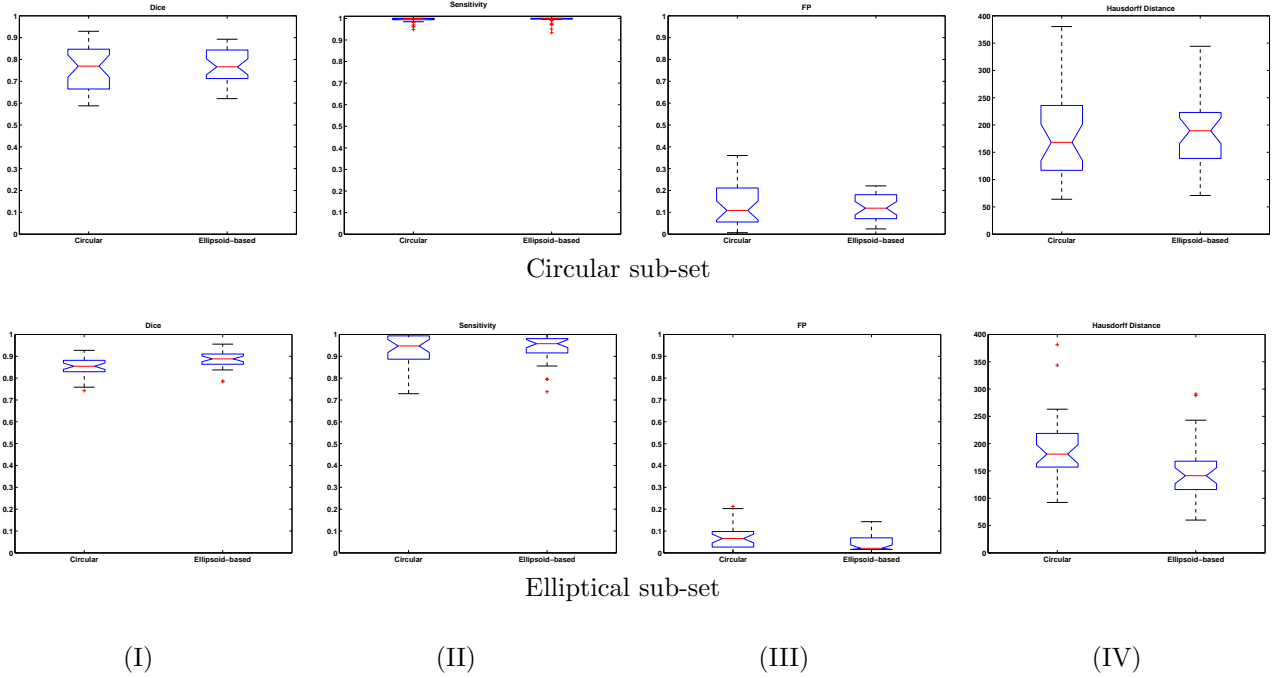


Figure 6. A comparison between circular (left) and ellipsoid-based (right) shape priors. Top row: Circular sub-set results. Bottom-row: elliptical sub-set results. Presented are box plot results for: (I) Dice, (II) Sensitivity, (III) FP and (IV) Hausdorff Distance.

In order to accurately reflect the distribution of the different results over the two sub-sets, notched box-and-whisker plots¹⁵ (hereon termed box-plots), are used throughout the different experiments. The boxes within the plot have lines at the lower quartile, median, and upper quartile values of the data. The whiskers are lines extending from each end of the boxes to show the extent of the rest of the data, this extent is set to $1.5IQR$, where IQR is the interquartile range of the samples. Outliers are data with values beyond the ends of the whiskers (marked as +). The notches represent a robust estimate of the uncertainty about the medians for box-to-box comparison. Boxes whose notches do not overlap indicate that the medians of the two groups differ at the 5% significance level. When comparing between different box plots, the following rules are set in order to specify the best results: 1) The box-plot should have the highest (lowest) median, significantly different from the other medians. 2) The interquartile range (IQR) should be relatively narrow. 3) The values of the lower and upper quartiles should be as high (low) as possible. 4) The amount of outliers should be low.

Figure 6 presents a comparison via box-plots, between the results of the circular (left) and ellipsoid-based (right), shape priors for each of the sub-sets. The results for the circular sub-set are presented in the top row and for the elliptical sub-set are presented in the bottom row. Columns (I)-(IV) show results for: Dice, Sensitivity,

False-Positive (FP) and Hausdorff distance, in respective order.

The box-plot results of the circular sub-set are similar for the two shape priors. The IQR obtained by the different quality measures for the ellipsoid-based prior is narrower than this obtained for the circular prior, but the median value is not significantly different between the two methods. This observation indicates that the ellipsoid-based prior can capture well a circular cervix region, as the ellipsoid is a generalization of a circle. In the box-plots obtained for the elliptical sub-set the ellipsoid-based prior obtained significantly better results, as expected.

4. CONCLUSIONS & FUTURE WORK

This work concentrates on embedding shape information for cervix boundary segmentation. Two methods are presented: the first method is based on a feature histogram distribution using a circular model. The second method is a novel method for embedding an ellipsoid-based model, which uses a distance measure between two level-sets. Experiments are conducted on two test-sets of cervigrams, and compared to human-expert markings. The incorporation of the prior shape information in the segmentation framework is shown to improve former data-driven segmentation results.

The results obtained with the two shape priors on the entire test-set are similar, as there are other factors that influence the curve evolution process. One of these factors is the initial, data-driven curve and its position relative to the true cervix boundary. This position varies considerably across the cervigram images due to their large content variability. The position of the initial curve has a strong influence on the active contour process, as the process tends to converge into local optimum.

A more focused experiment that isolates the effect of the shape-prior term on the segmentation quality, indicates the superiority of the ellipsoid-based prior over the circular prior. The ellipsoid-based prior (using the level set framework) is also shown to have a reduced order of complexity, $O(x)$, as compared to the complexity, $O(x^2)$, of the circular prior (and the corresponding curve-based functional).

The matching of the shape model to the problem at hand is an additional issue that should be considered. Although the manual ground-truth markings have a clear tenancy to resemble a circular or elliptical shape, this is not the case for all images. In part of the images, the segmentation quality when using the presented shape model, will suffer from lack of resemblance to the ground-truth. There are also cases where one prior overestimates the cervix region while the other prior underestimates it. Future work will include the combination of the two methods in order to produce better results. In addition, different ways for per-image parameter tuning will be investigated as a single set of parameters for the active contour framework might be inadequate for the current large and diverse database.

ACKNOWLEDGMENTS

We would like to thank the Communications Engineering Branch, NLM, NIH and the Hormonal and Reproductive Epidemiology Branch, NCI, NIH, for the data and support of the work.

REFERENCES

1. M. Schiffman and M. Adriaenza, "ACUS-LSIL triage study: design, methods and characteristics of trial participants," *Acta Cytol* **44**(5), pp. 726–742, 2000.
2. S. Gordon, G. Zimmerman, R. Long, S. Antani, J. Jeronimo, and H. Greenspan, "Content analysis of uterine cervix images: Initial steps towards content based indexing and retrieval of cervigrams," in *Proc. of SPIE Medical Imaging*, **6144**, pp. 1549–1556, 2006.
3. G. Zimmerman, S. Gordon, and H. Greenspan, "Automatic landmark detection in uterine cervix images for indexing in a content-retrieval system," in *Proc. of IEEE International Symposium on Biomedical Imaging*, pp. 1348–1351, 2006.
4. T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape modelstheir training and application," *Computer Vision and Image Understanding* **61**(1), pp. 38–59, 1995.

5. M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, **1**, pp. 316–323, 2000.
6. M. E. Leventon, O. Faugeras, W. E. L. Grimson, and W. M. Wells, "Level set based segmentation with intensity and curvature priors," in *Proc. of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '00)*, pp. 4–11, 2000.
7. Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision* **50**(3), pp. 315–328, 2002.
8. M. Rousson and N. Paragios, "Shape priors for level set representations," in *Proc. of 7th European Conference on Computer Vision-Part II (ECCV '02)*, pp. 78–92, 2002.
9. A. Litvin and W. C. Karl, "Coupled shape distribution-based segmentation of multiple objects," in *Proc. of Information Processing in Medical Imaging (IPMI)*, pp. 345–356, 2005.
10. R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distribution," *ACM Transactions on Graphics* **21**(4), pp. 807–832, 2002.
11. T. Riklin-Raviv, N. Kiryati, and N. Sochen, "Prior-based segmentation by projective registration and level sets," in *Proc. of IEEE International Conference on Computer Vision (ICCV'05)*, pp. 204–211, 2005.
12. T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Processing* **10**(2), pp. 266–277, 2001.
13. G. Zimmerman and H. Greenspan, "Automatic detection of specular reflections in uterine cervix images," in *Proc. of SPIE medical imaging*, **6144**, pp. 2037–2045, 2006.
14. G. Gerig, M. Jomier, and M. Chakos, "Valmet: A new validation tool for assessing and improving 3D object segmentation," in *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI'01)*, **2208**, pp. 516–523, 2001.
15. R. McGill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *The American Statistician* **32**(1), pp. 12–16, 1978.