Automatic Detection of Anatomical Landmarks in Uterine Cervix Images

7 authors, including:

Some of the authors of this publication are also working on these related projects:

Early detection of Osteoporosis and vertebral injury View project

Automatic x-ray screening for TB View project

Automatic Detection of Anatomical Landmarks in Uterine Cervix Images

Hayit Greenspan*, Shiri Gordon, Gali Zimmerman, Shelly Lotenberg, Jose Jeronimo, Sameer Antani, and Rodney Long

Abstract—The work focuses on a unique medical repository of digital cervicographic images ("Cervigrams") collected by the National Cancer Institute (NCI) in longitudinal multiyear studies. NCI, together with the National Library of Medicine (NLM), is developing a unique web-accessible database of the digitized cervix images to study the evolution of lesions related to cervical cancer. Tools are needed for automated analysis of the cervigram content to support cancer research. We present a multistage scheme for segmenting and labeling regions of anatomical interest within the cervigrams. In particular, we focus on the extraction of the cervix region and fine detection of the cervix boundary; specular reflection is eliminated as an important preprocessing step; in addition, the entrance to the endocervical canal (the "os"), is detected. Segmentation results are evaluated on three image sets of cervigrams that were manually labeled by NCI experts.

Index Terms—Cervical cancer, curvature features, image segmentation, landmark extraction, medical image analysis.

I. INTRODUCTION

THIS work is part of an ongoing effort towards the creation of a content-based image retrieval (CBIR) system for cervicographic images, otherwise termed "cervigrams." Cervigrams are currently being investigated as a means for detection, diagnosis and basic research of cervical cancer. Automated cervigram analysis tools are needed for these tasks.

*H. Greenspan is with the Department of Biomedical Engineering, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel (e-mail: hayit@eng.tau.ac.il).

S. Gordon was with the Department of Biomedical Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel. She is currently with Superfish Ltd. Givat-Shnuel, Israel (e-mail: shiri@superfish.com).

G. Zimmerman is with the Department of Biomedical Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel (e-mail: zimer@post.tau.ac.il).

S. Lotenberg was with the Department of Biomedical Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel. She is currently with Mobileye Vision Technologies Ltd. Jerusalem, 91450, Israel (e-mail: shelly.lotenberg@mobileye.com).

J. Jeronimo was with the Hormonal and Reproductive Epidemiology Branch , National Cancer Institute (NCI), National Institutes of Health, Bethesda, MD 20892 USA. He is currently with the Program for Appropriate Technology in Health (PATH), Seattle, WA 98107 USA (e-mail: jjeronimo@path.org).

S. Antani and R. Long are with the Communications Engineering Branch, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894 USA. (e-mail: santani@mail.nih.gov, rlong@mail.nih.gov).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

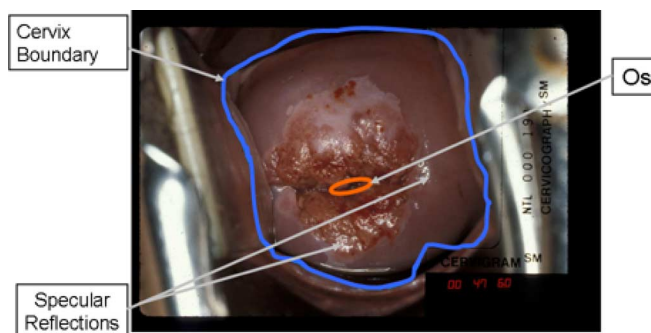Digital Object Identifier 10.1109/TMI.2008.2007823

Fig. 1. An example cervigram: the cervix boundary, the os, and SR artifacts are marked.

We start with an introduction to the clinical perspectives on cervicography and proceed to the image analysis challenges.

A. Cervicography in Cervical Cancer Research

Cervical cancer, the second most common cancer affecting women worldwide and the most common in developing countries, can be cured in almost all patients, if detected by high quality screening, and treated. However, cervical cancer incidence and mortality remain high in resource-poor regions, where high quality Pap (cytology) screening programs often cannot be maintained because of inherent complexity and cost. An alternative method of cervical cancer screening, called visual inspection with acetic acid, is based on color change of cervix tissues when exposed to acetic acid. This inexpensive method helps to detect abnormal cells that turn white (acetowhite) following the application of 3%–5% acetic acid [1]. An analogous photographic method that permits archive and study is cervicography. In this method the uterine cervix is photographed with a special fixed-focus 35-mm camera equipped with a ring flash that is used to provide enhanced illumination of the target region. Fig. 1 shows an example cervicographic image. For epidemiologic investigations, a cervigram resembles a low-magnification colposcopic image. When additional screening techniques are available, visual methods like cervicography may be used at the initial examination level, and patients with indicators of concern are then referred to colposcopic and/or Pap smear screening, or to treatment. Alternatively, DNA testing for a major risk factor for cervical cancer, human papillomavirus (HPV) infection, could be used as a primary screen, with visual triage.

The National Cancer Institute (NCI) has collected a substantial amount of biomedical information related to the

occurrence and evolution of uterine cervical cancer in longitudinal multiyear studies carried out in Guanacaste, Costa Rica, and in the United States. The Guanacaste Project is an intensive, seven-year population-based cohort study of human papillomavirus (HPV) infection and cervical neoplasia among 10 000 women in Guanacaste, where the rates of cervical cancer are perennially high. The ASCUS-LSIL Triage Study (ALTS), conducted in the United States, is a randomized clinical trial of management strategies for minor cervical cytological abnormalities, with two years of semi-annual follow-up. State-of-the-art visual, microscopic, and molecular screening tests are being used in these studies to examine the origins of cervical pre-cancer/cancer. The Guanacaste and ALTS projects now have a variety of subprojects based on collected specimens, visual images, and outcomes. NCI is examining several potentially important etiologic co-factors, which may contribute to cervical cancer risk. Most ambitiously, over 30 000 cervical cells and 30 000 plasma specimens are being tested for HPV DNA and serologic markers among 5000 women, to determine type-specific HPV DNA types and variants (there are over 40 types of cervical HPV). Data collected included patient age, sexual/reproductive history, laboratory test results; including Pap smear and cytology, and 100 000 cervicographic images in the form of 35-mm color slides, as well as medical classifications for the cervigrams into diagnostic categories [2]–[4]. NCI, NLM and the American Society for Colposcopy and Cervical Pathology (ASCCP) have formed the NIH-ASCCP Research Group that plans to use these images for the training and education of colposcopic practitioners [5]. A major long-term objective is to develop a unique web-accessible database of digitized cervix images for investigating the role of HPV in the development of cervical cancer and its intraepithelial precursor lesions in women. This work is a collaboration within the National Institutes of Health (NIH) between NCI and the National Library of Medicine (NLM) [6].

B. Automated Cervigram Analysis

The images within the NIH archive are unlabeled and have no attached annotation. Automated analysis of the cervigram images is thus needed in order to extract visual information from each individual image, across the large set of archived images. The extracted information can then be used to support cervical cancer research, to assist in training of experts and to enable future computerized cancer screening. In this work we present initial automated capabilities of landmark detection and extraction within the cervigram archive. We start by describing the individual cervigram image and describing the challenges of the landmark detection task.

A typical cervigram is presented in Fig. 1. The *cervix region* which is the main region of interest within the cervigram, is located in the central part of the image, with surrounding vaginal walls and parts of clinical equipment, such as speculum or swab, 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 defines the region of medical and anatomical interest within the cervigram and enables further analysis to

focus within the cervix region itself. An additional important landmark is the *os* which is the opening of the cervix. The shape, size and color of the os may vary strongly among cervigrams, but it is usually clearly visible. The os landmark is used by the medical experts as a reference point for interpreting cervix anatomy.

Automated analysis of cervigrams is a very complex and challenging task due to a variety of factors. First, the cervigram acquisition process involves the use of a strong camera flash in order to achieve good illumination of the convex shape of the cervix. Several artifacts are generated during this acquisition process, including a strong shading artifact that causes an inhomogeneous appearance within and across the tissues, as well as a specular reflection artifact that interferes with the automated analysis. Second, a large variability is present within the cervigram archive: the image acquisition setup is not constant. The viewing angle varies significantly across the images, causing the cervix region to differ in intensity and shape from one image to another. In addition, the physical scene that is imaged has intrinsic variability. For example, in different patients the cervix is not the same size, and additional noncervix tissues or medical instruments may exist. A third significant difficulty is the variability of cervix tissue content within the images, as not all defined tissue types are present in each cervigram image. Finally, the narrow dynamic range of colors and the lack of distinct boundaries between tissue regions, introduce additional, challenging image analysis and data classification tasks.

Initial studies can be found on the analysis of individual cervigram images, or the higher-resolution colposcopic images. Most of these studies are semi-automated, requiring the user to mark regions of interest on various cervix tissues [7]–[9]. Features such as color [9], texture [8], and shape [7] are then extracted. Based on these features the different regions are associated with different cervix tissues using various classifiers, such as neural networks [7] or the minimum distance classifier [8]. Additional works have started to address the task of fully-automated colposcopic image analysis [10], [11]. Preliminary segmentation efforts for the cervigrams within the NIH database were recently introduced [12]–[14]. The works to date usually focus on one specific analysis task (e.g., single landmark or tissue) or show initial results, with a small number of image examples. Several of the works published in this area, as well as companies working in related fields, are focusing on developing novel acquisition protocols to enforce a more controlled imaging environment, and to facilitate the image analysis tasks by utilizing very advanced imaging or fusion of imaging modalities. One such example is the LUMA imaging system [15], a commercial device that has recently been approved by the U.S. Food and Drug Administration as an adjunct to colposcopy. This system combines computer image analysis of video of the cervix, along with fluorescent light imaging and broadband white light imaging, in an integrated real-time algorithm for tissue classification.

The above approaches are not suited for handling an existing large image repository, such as the NIH cervigram archive. For the NIH archive, the analysis tools need to be robust enough to handle the large variability known to exist across images; it is highly desirable, however, that the tools be general enough to
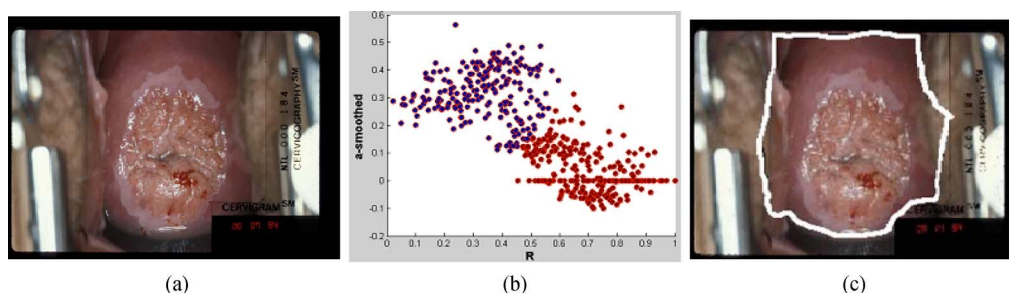
Fig. 2.   ROI detection. (a) Original cervigram. (b) Clustering in the selected feature space: a (smoothed) versus R. Darker gray pixels (purple in color version) are associated with the ROI. (c) Detected ROI outlined in white.

facilitate analysis across additional, similar archives of cervix images. These are the challenges of the current work; we focus on the tasks of cervix boundary segmentation, os detection, and the important preprocessing step of specular reflection elimination.

C. Proposed Framework

We present an automated, multistage procedure for cervigram segmentation and landmark extraction. Each stage targets a specific region or landmark within the cervigram. We identify and use features that we have found to be effective for each particular processing step. We use unsupervised modeling at various stages for feature clustering and classification. Using unsupervised clustering enables an adaptive, image-specific analysis, within the large image archive. In addition it eliminates the need for image normalization across the archive and for multiexpert labeling, which are critical for a supervised learning approach. In addition to statistical modeling, we use curve evolution tools, with novel energy functionals that we derived specifically for the cervigram data.

In a preprocessing stage (Section II), a coarse region-of-interest (ROI) is extracted from the given input image. The detected ROI is intended to exclude as much irrelevant information as possible, while making sure that the entire cervical area is included. A second important preprocessing step addresses the problem of specular reflection. Following the preprocessing stage, we are prepared to algorithmically delineate the cervix boundary (Section III). We propose a general model that incorporates edge, region, and shape information. The relevant edges in the image are described by the local curvature of the cervical surface. An algorithm for the detection of the os landmark is described next (Section IV). A geometric measure of local concavity of the gray-level image surface is used in conjunction with statistical modeling for the detection of the os region. Experiments and results are presented in Section V. A discussion concludes the work in Section VI. Parts of our work have been recently presented [12], [16], [17].

II. PREPROCESSING STAGE: EXTRACTING ROI AND REMOVING SPECULAR REFLECTION ARTIFACTS

The first two steps of the cervigram analysis framework are preprocessing steps essential to the successful computerized analysis of cervigrams. First, irrelevant image regions are discarded by an automated ROI detection algorithm. In the second step, regions with specular reflection (SR) are detected and

eliminated. Handling SR is important in a variety of medical applications, and an extensive overview can be found in [17]. Here we present the main algorithm and its results within the overall framework of our goals for processing of these cervix images.

A. Initial Automated ROI Detection

The cervix region is a relatively pink region located near the image center. For an initial delineation of the cervix, we use two features: the a color channel of the Lab color space (the higher the value of a, the "redder" the pixel color) and the distance R of a pixel from the image center. The R feature provides spatial information and supports the extraction of continuous regions within the image plane. The a color channel is initially smoothed in order to eliminate small details and the two features are normalized by their maximum value in order to obtain a similar range of values.

The image is separated next into two clusters in the 2-D ($a - R$) feature space; we use Gaussian mixture modeling, initialized by a K-means procedure, as a statistical clustering methodology [18]. The cluster that has the highest a-mean and the lowest R-mean is selected. The ROI is chosen as the largest connected component within the pixels associated with this cluster. Postprocessing of the ROI includes morphological boundary smoothing and elimination of small holes. Morphological operations are performed with a small structuring element of a fixed size. An example of the clustering process and final ROI is presented in Fig. 2.

Criteria for judging an acceptable ROI are that it should exclude irrelevant information such as medical equipment, frames, text, and noncervix tissues (Fig. 1), while preserving the cervix region in its entirety. These requirements are largely satisfied in all the cervigrams analyzed in the current work. After the ROI has been detected, all subsequent processing takes place in the ROI interior; hence, confusing patterns and textures in the cervigram periphery will be ignored and will not degrade algorithm performance.

B. Detection and Elimination of SR

SR, or highlights, are small and bright regions on the cervix surface, which are generated during the image acquisition process by strong reflectors, such as fluids (Fig. 1), on the surface of the cervix. These SR artifacts interfere with the content analysis of the regions surrounding them. The bright white regions of SR may be confused with lesions, which are
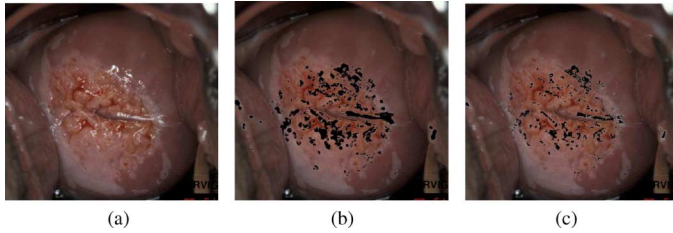
Fig. 3. Identification of SR. (a) Original cervigram (cropped around the ROI). (b) Candidate SR regions (black). (c) Final identification of SR (black).

usually relatively bright. Furthermore, the strong gradients created by SR amplify the local contrast, causing potentially erroneous results in any further processing that incorporates gradient or contrast measures. Reliable identification of SR is therefore essential.

We present a description of the proposed algorithm for SR detection and removal from the cervigram images (see also [17])[1]. Modeling of SR pixels within an image using statistical tools is a difficult task, since the number of pixels belonging to SR regions is extremely small relative to the image size. SR pixels tend to be associated with other clusters, rather then being recognized as a separate cluster. To address this problem, we propose a two-stage process in order to model the unbalanced clusters. In the first stage, candidate SR regions are detected by using thresholds of high brightness I and low color saturation S as suggested in [19]. In the second stage, the candidate SR mask is further refined by selecting only the pixels in the vicinity of strong gradients. Pixels within the candidate regions form a subset of the entire image pixel set, as shown in Fig. 3(b). Within this reduced set, the amount of SR and non-SR pixels is more balanced, and the two modes can be statistically modeled as a mixture of Gaussians.

Our experiments suggest that SR pixels are optimally modeled by two Gaussians. One represents the brightest, almost white pixels, that have lost most of their chromatic information. The second Gaussian represents the less bright pixels that retain some of their original color. The non-SR regions are modeled by two additional Gaussians to account for various tissue types. Thus, an overall mixture of four Gaussians is used to model the candidate regions. Following the modeling stage, pixels corresponding to the two Gaussians with the highest mean intensities are selected as the SR pixels. Fig. 3(c) shows the final set of SR labeled pixels following the statistical modeling and segmentation.

In the processing steps that follow, the SR pixels are usually ignored. Note that for some purposes it may be desirable to fill in the "holes," i.e., to introduce new pixel values for the black pixels in Fig. 3. For example, a filled-in image may also be of value to the clinician for diagnostic purposes, to eliminate the visual distraction of bright SR; in this case, it would likely be desirable to also display the original image and/or the image with SR removed, but not filled in, so that the artificially-created regions would not be confused with real tissue. We use a filling scheme that eliminates the strong gradients associated with the SR, while preserving the original texture. In this scheme the

average color of the surrounding pixels is propagated into the specular regions, creating a smooth filling of the image. This approach is based on the observation that the highlights (the SR) formed on the moist surface of the cervix are very small. The color underneath the highlights is assumed to be nearly constant and similar to the color of the pixels in the immediate surroundings. The SR detection and elimination process is evaluated in [17]. Image examples are presented in Section V-C.

The output of the preprocessing phase of the proposed framework is a smaller-size image that contains the cervix region and a set of labeled SR pixels. The following processing steps analyze this generated image further, towards the accurate delineation of landmarks of interest.

III. CERVIX BOUNDARY DETECTION

The ROI extracted in the preprocessing step is coarse and often includes large parts of the vaginal region. An additional step is required in order to refine the detection quality of the cervix boundary. Various segmentation methods can be used for this purpose, including region growing and active contours. In region growing [20] a region is defined via propagation of similar neighboring pixels. Two scenarios using region growing can be defined in the cervix boundary detection task. In the first, the region is propagated from a single point located in the center of the initial ROI. The selection of adequate features for the propagation is not an easy task, as different images will require a different set of features in order to advance the region from the center of the cervix to its boundaries. The region growing can be disrupted by other tissues within the cervix. In a second scenario, multiple points on the boundary of the initial ROI can be used as seed points and propagated inwards. The boundaries generated in this way cannot be restricted by any smoothing or shape constraints.

Segmentation that combines edge and region information can be achieved using an energy minimization techniques via the active contour framework. This framework can be subcategorized into snakes [21] and level set [22] methods, two different schemes to carry out the contour deformation process. A review and comparison between different energy functionals can be found in [23]. A main conclusion of the review, in which both methods were evaluated on a set of different medical images, was that the integration of forces from different energy functionals may lead to better segmentation results. The main advantages of such methods, as compared to region growing, are their ability to integrate local and global information and to account for both region and edge features, while preserving smoothness of the boundaries.

In the current work energy minimization via active contours is used in order to refine the initial ROI so that it matches the actual cervix boundaries more closely. The main contribution of the current work is the energy functional used, that consists of forces and features adequate for the task of cervix boundary detection. Region, edge, and prior shape information are all used for this purpose. We use the implicit implementation via level sets. Implementation via the parametric snake mechanism may be possible as well.

In an active contour framework the image is considered as a function $I : \Omega \rightarrow \Re^+$ where $\Omega \in \Re^2$ is the image domain.

---

[1]Available online at http://www.eng.tau.ac.il/~hayit/

The segmentation problem is mathematically formulated as the search for a contour $C : [0, L] \rightarrow \Re^2$ in the image, which is optimal with respect to some predefined integral measure, $E(C)$, also called the energy functional. Formally, this problem is stated as: $C = \arg \min_C E(C)$. In our work the energy functional consists of two terms: a data term, $E_{\text{data}}(C)$ and a shape prior term, $E_{prior}(C)$. In the following sections, we present a full description of the curve evolution process. The data term is described in Section III-A. The prior shape term is described in Section III-B. The combination of the two terms is discussed in Section III-C.

A. Curve Evolution Based on Image Data

The data term is based on the following general integral measure: $E_{\text{data}}(C) = \int_C g(C(s))ds + \iint_{\Omega_C} f(x, y)dxdy$, which imposes constraints on the contour, as well as on region properties inside and outside the contour.

The extremals of the chosen energy functional are identified by the Euler-Lagrange equation: $\delta E(C)/\delta C = 0$ and are found by gradient descent in a level set implementation [22]. In this level set formulation a closed curve C is represented implicitly by embedding in a higher dimensional function $\phi(x, y)$, where C is its zero set, $C = \{\{x, y\} : \phi(x, y) = 0\}$.

The data term in the current work is based on a curve evolution functional that incorporates edge and region-based information, as suggested for fast edge integration by Kimmel [24]. The compact representation of the level set formulation is the following curve evolution equation:

$$\phi_t^{\text{data}} = \left[f(x, y) + \text{sign}(\langle \nabla \phi, \nabla I \rangle)\Delta I \right.$$
$$\left. + \alpha \, \text{div} \left(g(x, y) \frac{\nabla \phi}{|\nabla \phi|} \right) + \beta(c_2 - c_1)\left(I - \frac{c_1 + c_2}{2} \right) \right] |\nabla \phi| \quad (1)$$

where I is a gray level function of the image and ϕ is the level set function.

The first term of (1) is the weighted region term, which advances the curve according to a scalar weight function, $f(x, y)$. The second, robust alignment term, influences the solution curve to align with edges within the image. The third term, which we refer to as the GAC term, is derived from the theory of Geodesic Active Contours. It attains low values for the portions of the contour that overlap image edges, thus preventing these portions from further evolution. In this term, the function $g(x, y)$ is an inverse edge indicator, usually of the form $g(x, y) = 1/\sqrt{1 + |\nabla I|^2}$. The fourth and final term is the minimal variance term, which attempts to separate the foreground and the background of the image with respect to their relative mean values. The two constants, c_1, c_2, are calculated as the mean intensities in the interior and the exterior of the contour, respectively.

A direct application of the level set formulation of (1) for the detection of cervix boundaries was experimentally found to be ineffective; we attributed this to the presence of irrelevant edges within the cervigram that are formed by folds of surface tissue and by the various tissues within the cervix. These edges interfere with the curve attraction to the cervix boundaries. For

this reason, we have modified the original formulation of (1) to adapt to the unique characteristics of the cervigrams. We note again here the importance of the initial preprocessing to crop the image to the coarse ROI. This ensures that the strong edges of the image frame and of the medical instruments do not attract the evolving curve. This curve C is initialized as the boundary of the coarse ROI, which is always larger than the desired final contour. The weighted region term (term 1) is defined as $f = -1/t$, with t being the iteration number, so that it will have decreasing influence over time, i.e., as the number of iterations increase. This force is set negative to ensure that the contour moves inward from its initial state. The minimal variance term (term 4) uses the a color channel (of the $CIE - Lab$ color space), to represent the cervix region's pink coloring. It is assigned a very low weight to reflect the fact that the color difference between the interior and exterior of the cervix is usually not significant.

The multitude of irrelevant edges in the cervigram image makes the gradient-based terms (robust alignment and GAC, terms 2 and 3) inappropriate for the task of cervix boundary detection. We propose an alternative edge indicator, based on the cervix *convexity*. Using convexity as a characterizing feature is motivated by the observation that most of the cervix boundaries are outlined by folds of surface tissue that form narrow valleys and are distinctively concave. The boundaries are easily detected by their largest positive principle curvature, $k1$. Edges generated by the color transition between two different tissue types are expected to have strong intensity gradients but low curvature; thus, the presence of different tissue types is not expected to interfere with the curve evolution for cervix boundary detection.

The principle curvatures measure the maximum and minimum bending of the image surface, $I(x, y)$, at each point [25]. The two principal curvatures and directions are obtained as the eigenvalues and eigenvectors, respectively, of the following matrix:

$$\frac{1}{g_{11}g_{22} - g_{12}^2} \begin{bmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix} \quad (2)$$

where g_{ij} and b_{ij} are related to the first and the second fundamental forms, respectively, and are given by

$$\{g_{11}; g_{12}; g_{22}\} = \{I_x^2 + 1; I_x I_y; I_y^2 + 1\} \quad (3)$$
$$\{b_{11}; b_{12}; b_{22}\} = \frac{1}{\sqrt{1 + I_x^2 + I_y^2}}\{I_{xx}; I_{xy}; I_{yy}\}. \quad (4)$$

The largest principle curvature, $k1$, and its direction $\vec{V}$ are next used for the generation of the curvature-based vector field of the new edge indicator. The $k1$ feature describes both concave (inward bending) and convex (outward bending) regions, corresponding to positive and negative values, respectively. The cervix boundaries are concave; therefore, we adapt the $k1$ feature to favor detection of only concave regions. Specifically, we normalize $k1$ to lie in the range between 0 and 1, such that concave regions will have $k1$ near 1, and convex regions will have $k1$ near 0. This adapted $k1$ is defined by

$$\tilde{k_1} = \frac{(k_1 - \min(k_1))}{(\max(k_1) - \min(k_1))}. \quad (5)$$
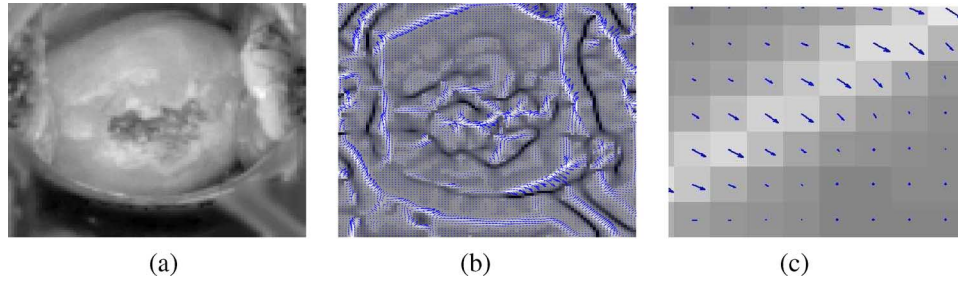
Fig. 4. Curvature-based boundary indicators. (a) Smoothed intensity image of a cropped cervigram. (b) Scaled principle directions overlaid on a map of the normalized principle curvature. Concave regions—bright. Convex regions—dark. (c) Zoom-in on one of the strong edges in (b).

The unit vector of the principle direction $\vec{V}$ is scaled by the associated $\tilde{k}1$, thus emphasizing directions of concave regions and suppressing directions of convex regions. The curvature measures are computed on a smoothed version of the input cervigram in order to eliminate small-scale variations. The resulting boundary indicators are illustrated in Fig. 4. In Fig. 4(b) and (c) the scaled principle directions are overlaid on a map of the normalized principle curvatures. A correlation can be seen between regions with strong $\tilde{k}1$ values (bright) and small $\tilde{k}1$ values (dark), with concave and convex cervigram regions, respectively.

We propose to use the normalized curvature feature, $\tilde{k}1$, and its associated principle directions, $\vec{V}$, to modify the gradient-based terms (robust alignment and GAC) of (1). The GAC inverse edge indicator function g is replaced by

$$\tilde{g}(x,y) = \frac{1}{\sqrt{1+\tilde{k_1}^2}}. \qquad (6)$$

The robust alignment term is replaced so that it is driven by the vector field of the principle curvature direction $\vec{V}$, scaled by the normalized curvature. The modified alignment term is

$$\text{sign}\left(\langle \nabla\phi, \tilde{k}_1\vec{V}\rangle\right)\text{div}(\tilde{k}_1\vec{V})|\nabla\phi|. \qquad (7)$$

The new curvature-based level set formulation of the data term is

$$\phi_t^{\text{data}} = \left[-\frac{\gamma}{t} + \eta\text{sign}\left(\langle\nabla\phi,\tilde{k}_1\vec{V}\rangle\right)\text{div}\left(\tilde{k}_1\vec{V}\right)\right.$$
$$\left. +\nu\text{div}\left(\tilde{g}(x,y)\frac{\nabla\phi}{|\nabla\phi|}\right) + \beta(c_2-c_1)\left(a - \frac{c_1+c_2}{2}\right)\right]|\nabla\phi| \qquad (8)$$

where γ, η, ν, β are the energy functional tuning parameters and a is the a-color channel of the $CIE-Lab$ color space.

Fig. 5 shows the advantage of using curvature features over the more commonly used intensity gradients. Two examples (I and II) are presented. The input images with the ground truth regions marked by experts are shown in (a). The manually marked cervix boundary, that serves as our ground truth, is shown as the surrounding curve (yellow). The automatic cervix boundary, found with the curvature-based functional, is shown on the original image as the inner curve (red), as shown in (b).
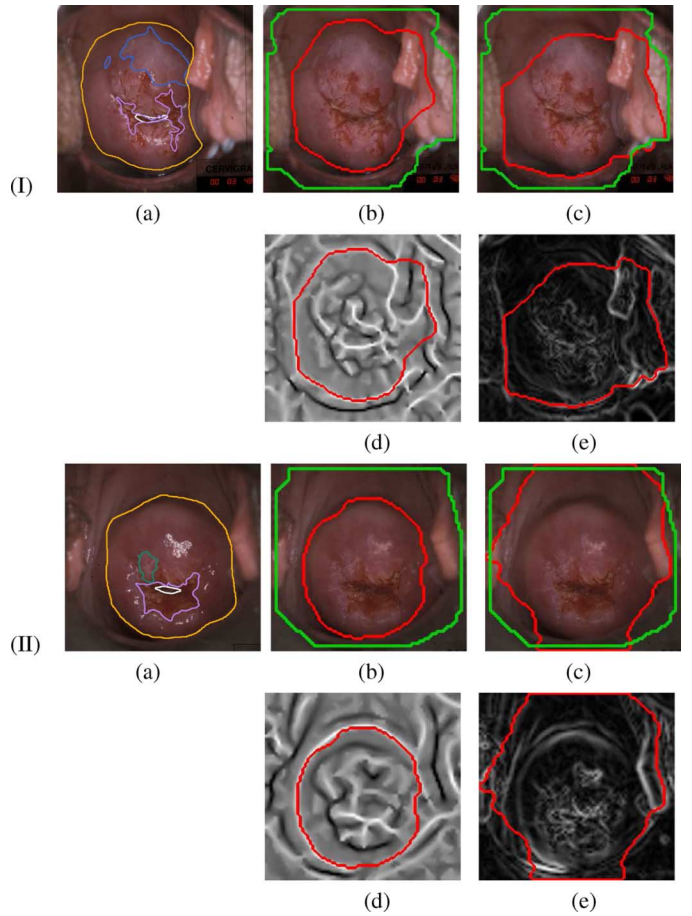

Fig. 5. Two cervigram examples (I) and (II). (a) Original cervigrams with tissues of interest marked. (b) Initial coarse ROI boundary (outer curve—green). Cervix boundary found using the curvature-based edge indicator (inner curve—red). (c) Cervix boundary found using the intensity gradient edge indicator (inner curve—red). Initial coarse ROI boundary (outer curve—green). (d) Curvature-based boundary contour (red) overlayed on the principal curvature function. Edges with greater relevance attain high values (white). (e) Intensity gradient function overlayed with the gradient-based boundary contour (red).

The same curvature based contour, marked on the principal curvature function, is presented in (d). The automatic cervix boundary, found with the gradient-based functional, is found as the inner curve (red) shown on the original image, in (c). The same gradient-based contour, marked on the intensity gradient, is presented in (e). The boundary of the coarse ROI, that serves as the initial condition for curve evolution, is the outer curve [green in (b) and (c)].

In example I, the AW lesion (blue) causes a color transition between the tissues. This transition introduces an intensity gradient (e), which attracts the solution curve in (c). The curvature feature (d) does not exhibit positive response to this particular edge, and thus the resulting curve in (b) remains aligned to the cervix boundary which is well defined by a prominent response in the curvature function.

In example II, the cervix boundary is defined by valleys that yield a very strong positive response in the curvature function (d). The curvature-based contour closely follows this response, resulting in an accurate cervix segmentation (b). The gradient based curve, on the other hand, is easily distracted by the regions that have stronger gradient magnitudes then the actual cervix boundary (c).

It is important to note that a curve evolution process is governed by all of the terms in the defined energy functional [eq. (8)]. Two of the terms are closely related to the concavity measure. When no strong concavity response is present in the vicinity of the curve, it is influenced by the minimal variance and the balloon terms. This results in a boundary that mostly, but not entirely follows the concave valleys.

B. Curve Evolution Based on a Prior Shape Model

Using the data term of the energy functional [eq. (8)] the detected cervix boundary is optimal with respect to color and curvature-based edge features. When observing human expert markings, an evident commonality is that the cervix boundary is marked as smooth and circular. The curve evolution framework is next adapted in order to accommodate this prior shape information and further refine the detection quality of the cervix boundary.

Applying prior shape information to image segmentation is a well studied problem. One popular methodology uses deformable models learned from a training set of registered images [26]–[28]. Such a training set is not available in the current case, so we instead incorporate circular shape information and do not require cross-image modeling or registration. We rely on a method based on the distribution of a shape feature, as given in [29]. We incorporate a shape-based energy term into the curve evolution model that penalizes the difference between the feature distribution of a given curve and the feature distribution of a prior reference curve. It can be shown that such distributions capture the intuitive similarity of shapes in a flexible way, while being invariant to shape transformations [30]. We start with a brief description of the method as presented in [29]. 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

$$H(\lambda) = \frac{\int_C h\{\Phi(C) < \lambda\}dw}{\int_C dw}. \quad (9)$$

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
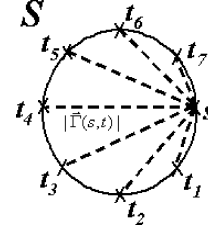

Fig. 6. Internode distances between points within the sampled set S, on the prior circular model.

will write $H(C, \lambda)$. The shape-based energy term, $E_{\text{prior}}(C)$, is defined as

$$E_{\text{prior}}(C) = \int_\Lambda [H^*(\lambda) - H(C,\lambda)]^2 d\lambda \quad (10)$$

where the prior shape information is captured in the target distribution $H^*(\lambda)$.

The shape descriptor that we use to describe the circular shape prior is termed the "internode 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 internode distances feature is [29]

$$\nabla E(\Gamma)(s) = 2 \int_{t \in S} n(s) \cdot \frac{\vec{\Gamma}(s,t)}{|\vec{\Gamma}(s,t)|}[H^*(F_{st}) - H(\Gamma, F_{st})]dt \quad (11)$$

where Γ is the parameterized curve as a function of the arclength $\{X(s), Y(s)\}$ 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 internode distance between nodes (s,t), is defined as

$$F_{st} = \frac{|\vec{\Gamma}(s,t)|}{\text{mean}(\{|\vec{\Gamma}(s,t)| \mid (s,t) \in S\})}. \quad (12)$$

Fig. 6 illustrates the internode distances between a single point, s, and the rest of the points, $t_1, \ldots, t_7$, within a sampled set S (S includes only eight points for illustration purposes). The histogram of the internode shape descriptor, $h(F_{st})$, is constructed out of all the available (normalized) distances between such node pairs.

We implement the level set formulation for the shape term as follows:

$$\phi_t^{\text{shape}} = \nabla E(\Gamma)(s) * G_\sigma(x,y) \quad (13)$$

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 eq. (13)] of size 5×5 and $\sigma = 0.5$. The sampled set S in our implementation includes all the points along the zero level set of the evolving curve.

C. Combination of the Data and Shape Terms in the Curve Evolution Process

The data and shape terms, Φ_t^{data} and Φ_t^{shape}, are combined in a two-stage procedure. The first stage uses the data term to evolve the initial, approximate ROI to better fit the cervix
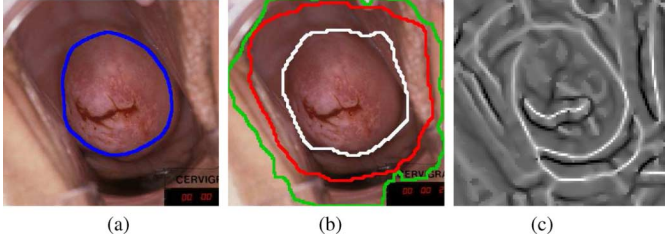
Fig. 7. Incorporating prior shape information in the curve evolution functional. (a) Manual marking of the expert (blue). (b) Cervix boundary results: initial ROI (green). Boundary detected using data and shape terms in a single-stage procedure (red). Boundary detected using the two-stage procedure (white). (c) Curvature feature map.

region. We will call the curve that is output from this stage 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 Fig. 7. Fig. 7(a) shows the medical expert-marked boundary (blue). Fig. 7(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 Fig. 7(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.

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 predefined 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:

$$\phi_t = \phi_t^{\text{data}} + w_d(x,y)\phi_t^{\text{shape}} \tag{14}$$

where 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

$$w_d(x,y) = \begin{cases} \frac{N}{|d(x,y)|}, & \text{if } d(x,y) < T \\ N, & \text{otherwise} \end{cases} \tag{15}$$

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. 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.

Fig. 8 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 model.

IV. Os Detection

The opening to the neck of the uterus is called the external os (here simply the os). The size and shape of the os varies widely with age, hormonal state, and whether the woman has had a vaginal birth. In women who have not had a vaginal birth the os appears as a small, circular opening. In women who have had a vaginal birth, the os appears wider, more elongated, and irregularly shaped. The os is an important anatomical landmark within the cervix region.

The detection of the os is not straightforward, since its shape, color and relative location in the image varies significantly across the images. Furthermore, its color is often similar to the color of the surrounding tissue (Fig. 9). However, one characteristic is consistent: the os region is always concave. Based on this observation, we introduce a geometric measure of local concavity for detection of the os. Concavity is related to the 3-D shape of an object. Since 3-D data of the cervix is unavailable in this image database, we use as a surrogate a geometric measure of local concavity for the gray-level image surface (introduced by Tankus *et al.* [31]). The local concavity operator that we use produces an image, termed D_{arg}, which has a strong response at concave or convex regions. Based on the D_{arg} image, we propose the following os detection algorithm: 1) compute the D_{arg} image; 2) detect high-intensity regions within D_{arg} (concave regions); 3) select as the os, the largest region that is close to the cervix center. We next elaborate on the proposed scheme, focusing on the computation of the D_{arg} image.

The D_{arg} detects surface patches that are locally concave or convex by detecting the ray of discontinuity in the gradient argument, $\theta(x,y)$, of the image function, $I(x,y)$. We describe below the details of the computation of the D_{arg} image (following [31]).
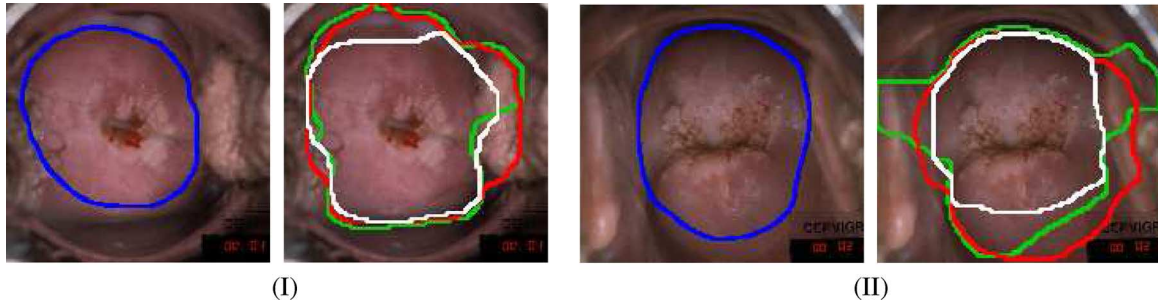
Fig. 8. Local weights effect on boundary detection quality for two cervigrams examples (I)–(II). Left image—manual markings of the expert (blue). Right image—boundary detection results. Data-driven curve (green). Equally weighted shape term (red). Locally weighted shape term (white).
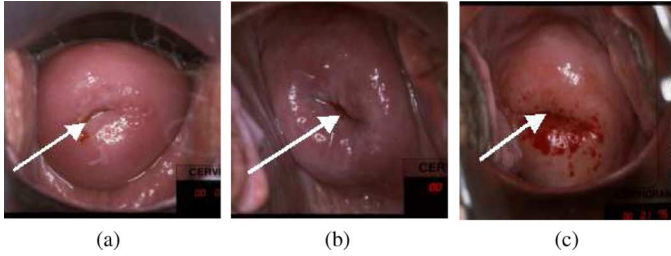

(a) (b) (c)

Fig. 9. Examples of the os in different cervigrams.

The gradient argument, $\theta(x,y)$, is defined as

$$\theta(x,y) = \arctan\left(\frac{\partial}{\partial y}I(x,y), \frac{\partial}{\partial x}I(x,y)\right) \quad (16)$$

where the 2-D $\arctan$ function is computed by

$$\arctan(y,x) = \begin{cases} \arctan\left(\frac{y}{x}\right), & \text{if } x \geq 0 \\ \arctan\left(\frac{y}{x}\right) + \pi, & \text{if } x < 0, y \geq 0 \\ \arctan\left(\frac{y}{x}\right) - \pi, & \text{if } x < 0, y < 0 \end{cases} \quad (17)$$

In order to obtain D_{arg}, a local convexity operator Y_{arg} is defined first. Y_{arg} is the derivative of the gradient argument in the y direction

$$Y_{\mathrm{arg}} = \frac{\partial}{\partial y}\theta(x,y). \quad (18)$$

Fig. 10 demonstrates how, for a locally concave paraboloid structure [Fig. 10(a)], the gradient argument has a strong discontinuity on the negative x axis [Fig. 10(c)]. Note that the coordinate system is placed in the center of the paraboloid, parallel to its axes. This structure can be easily detected by the convexity operator, Y_{arg} [Fig. 10(d)].

The convexity operator will have a strong response (in theory, infinite) on the negative part of a horizontal axis of a convex or concave patch. Its response to linear gradients or abrupt changes in the image are much weaker than the response to convex patches. Note that the proposed operator is insufficient for the purpose of cervix boundary detection (Section III), since it does not provide the direction and magnitude of principle curvature directions. However its stability and robustness make it very suitable for the os detection task.

The convexity operator may be extended in order to respond to other axes of the paraboloid [31]. In the current work, we
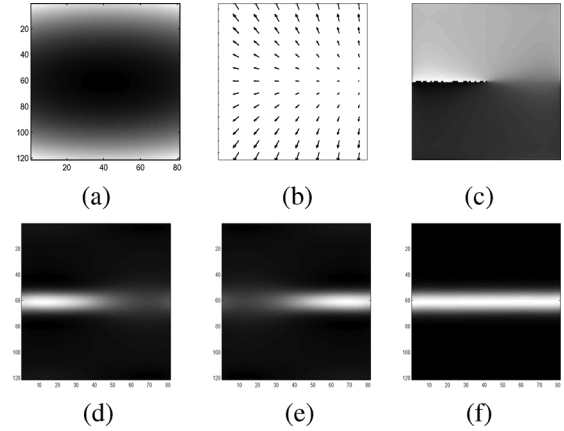

Fig. 10. (a) Synthetic concave paraboloid. (b) Intensity gradients, zoomed on the negative x-axis. (c) Gradient argument, $\theta(x,y)$. (d) Y_{arg}. (e) Rotated version of Y_{arg}—positive x-axis. (f) D_{arg}.
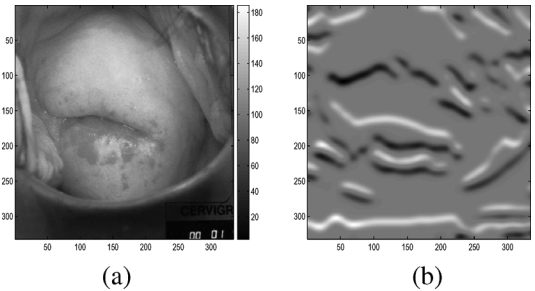

(a) (b)

Fig. 11. (a) An example cervigram. (b) Resultant D_{arg} image: the bright regions (high response values) correspond to concave regions within the cervigram.

are interested in detecting the os region, which is circular or nearly horizontal, thus a strong response along the entire x-axis is desired. This can be achieved with the D_{arg} image, which is an extension of Y_{arg}. The D_{arg} image is obtained by rotating the original image, I, by π degrees, calculating the Y_{arg} and then rotating the result image back [Fig. 10(e)]. The received response is then summed with the original Y_{arg} [Fig. 10(f)].

The D_{arg} image for an example cervigram is illustrated in Fig. 11. It can be seen that there is a strong positive response (light gray) for regions that look like concave horizontal cylinders. For the case of horizontal convex patches, the response is negative (dark gray).
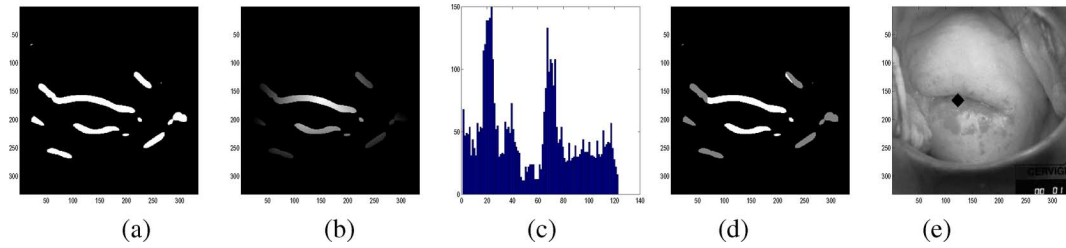
Fig. 12. Os detection process. (a) Mask of concave regions. (b) Distances of the mask pixels from the cervix boundary. White color denotes the highest distance value. (c) Histogram of the distance feature. x-axis: distance in pixels; y-axis: number of pixels per distance, within the mask. (d) Labeled mask pixels. The grey pixels were classified as close to the boundary, the white pixels as close to the center. (e) The cervix image with the center of the os marked by the black diamond.

TABLE I
SUBSETS WITHIN THE CERVIGRAMS ARCHIVE

	Set_1	Set_2	Set_3
# images	120	158	100
# experts	1	1	2
marked regions	cervix os	cervix os	cervix

TABLE II
QUANTITATIVE RESULTS FOR BOUNDARY DETECTION: CLUSTERING-BASED SCHEMES [14] VERSUS THE CURVE-EVOLUTION PROCEDURE. AVERAGE RESULTS ARE PRESENTED (Set_1)

	clustering	curve evolution
Dice	0.41	0.81
Sensitivity	1	0.94
FP	0.35	0.13

Multiple regions within the image achieve a high value of D_{arg}. Most of these regions are located on the horizontal boundaries of the cervix and need to be discarded. The os detection scheme is described next and illustrated in Fig. 12. Initially the D_{arg} response is automatically thresholded to generate a mask of candidate concave os regions [Fig. 12(a)]. Only the regions inside the previously detected cervix boundary are considered. Next, the distances between the pixels within the mask and the cervix boundary are computed by means of the distance transform [Fig. 12(b)—dark values correspond to small distances]. It can be seen from the histogram of the distance feature [Fig. 12(c)], that the pixels are naturally divided into two groups, one close to the cervix boundary and one close to its center. The pixels are clustered into these two groups by the K-means algorithm, producing the result in Fig. 12(d), where two values are shown: grey represents pixels close to the boundary and white represents pixels close to the center. The largest segment that has a majority of pixels associated with the center cluster is selected and the center of this segment is marked as the os landmark [diamond in Fig. 12(e)]. Os detection examples are shown and evaluated in Section V-B.

V. EXPERIMENTS AND RESULTS

Cervigram images were collected in the form of 35-mm color slides which were then digitized at 1660 dpi, compressed with 40:1 JPEG. Boundaries were marked and labeled with the NLM Boundary Marking Tool [32]. Three image sets are used in the experiments, as summarized in Table I. In Set_1 and Set_2 the boundaries for the cervix and the os were marked by a single medical expert with specialized experience in gynecological oncology. Set_3 includes cervigrams in which two experts (not necessarily the same experts across all of the images) have marked the cervix boundary. Qualitative and quantitative results of cervix boundary detection and os delineation are provided next.

A. Cervix Boundary Detection Results

We start by comparing the performance of the curve evolution process as presented in this work to clustering-based schemes for cervix-region extraction. One such scheme, which performs clustering in the $a - R$ feature space, was used in the initial ROI detection (Section II-A). Several variations on the features used, feature normalization, and clustering techniques, were recently quantitatively compared in [14], using the same NIH cervigram test data (Set_1). The best clustering-based overlap results from [14] are presented in Table II and compared to the average curve-evolution overlap results of the current work. In the quantitative evaluation generated contours are compared to expert-based markings, using multiple performance measures. Specifically, several overlap measures were used: the Dice measure: $2|S \cap R|/(|S| + |R|)$, the Sensitivity measure: $|S \cap R|/|R|$, and the False Positives (FP) measure: $|S \cap \hat{R}|/|\hat{R}|$, 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 results in Table II show a considerable improvement in the Dice and the FP measures using the curve evolution process. Sensitivity results slightly decreased, remaining at a strong level. These results indicate a more accurate delineation of the cervix region by the curve evolution process. Nonrelevant tissues, outside of the cervix region, are better defined, which is an important step for further cervix analysis.

The effectiveness of the prior shape model for refinement of the initial data-driven segmentation in the curve-evolution process, is explored next. In this experiment both Set_1 and Set_2 are used. A set of five cervigrams, randomly selected from Set_1, are used for the selection of algorithm parameters. In the first step of the curve evolution process we set the parameters of (8) to: $\gamma = -3; \eta = -20, \nu = 20; \beta = 0.01$. The time step was set to: $\Delta t = 2.5$. The addition of the shape term in the second step requires a modification of the weights of the curvature-based forces within the data term to: $\eta = -10; \nu = 10$. Number of iterations in the second step were limited to 20. We found that
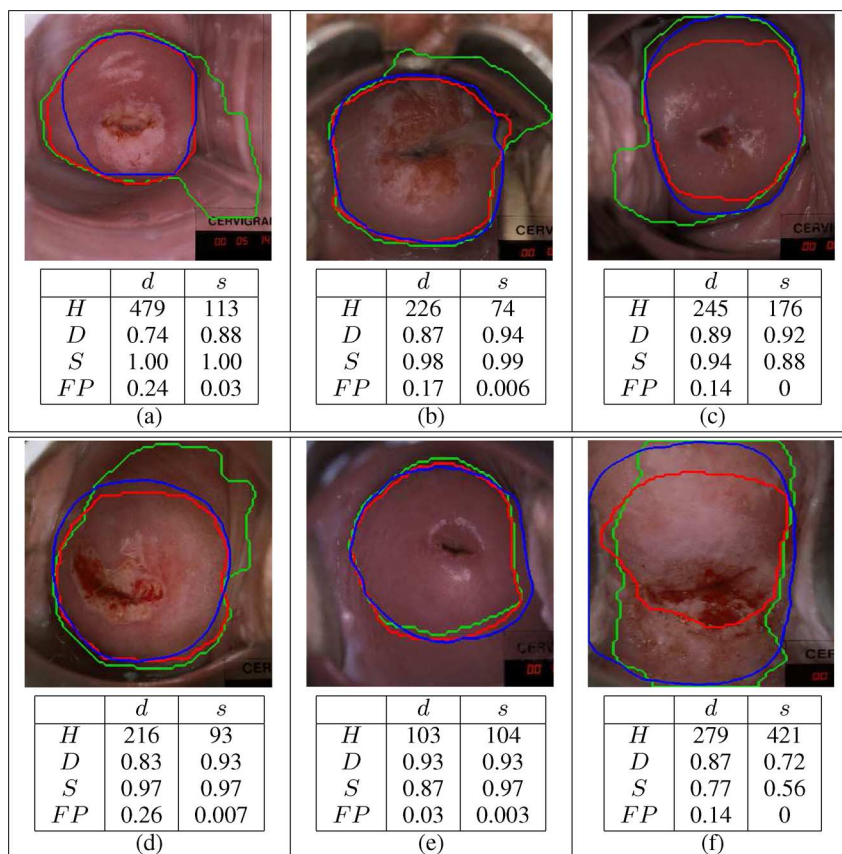
Fig. 13. Cervix boundary detection. Contour based on color and curvature features—marked in green. Final contour, following the shape-based refinement—marked in red; Expert markings imposed in blue. Hausdorff (H), Dice (D), Sensitivity (S), and False Positives (FP), for the data-driven (d) and the shape-based (s) contours, are listed under corresponding cervigrams.

these modifications were necessary to prevent the curve from shrinking into the cervix region in cases where its initial position is located within smooth regions and is already close to the desired result. The N parameter in (15) was set to $N = 4$. It is important to note that once defined, the parameters were kept constant across all images in the test set.

Fig. 13 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 prior shape model, is marked in red. Manual markings of the expert are shown in blue. Quantitative results are listed under each image example, using the overlap measure, along with the Hausdorff distance [33], which is a maximum surface distance. The Hausdorff distance defines the largest difference between two contours (measured in pixels) and is a good indicator for shape resemblance. A smaller Hausdorff distance corresponds to more similar contours. For the Hausdorff results, note that the images are of approximately 1500×2500 pixels.

Several observations can be made. 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 measures is achieved. The sensitivity results of examples (a)-(e) attain high values, as they include most of the expert's markings. These sen-

sitivity results are similar for the two contours being compared. Example (f) shows an exceptional case, where the sensitivity decreases due to the shape prior constraint. In this example the data-driven curve is already close to, 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 III summarizes average results of cervix boundary detection for Set_1 and Set_2 images, with and without the prior shape model. The Dice measure shows a small improvement and the 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 distance shows a substantial improvement when adding the shape term, thus indicating a better shape resemblance of the generated contours to the expert's markings. The FP measure shows a considerable improvement, corresponding to a strong reduction of nonrelevant tissues. This reinforces the addition of the shape term to the curve evolution process. Similar results are achieved for both sets.

In the final experiment the results of the suggested framework were compared to the markings of two experts, as available in Set_3. Fig. 14 presents several example segmentations. The manual markings of the two experts are shown in green and blue. The results of the algorithm are imposed in red. The

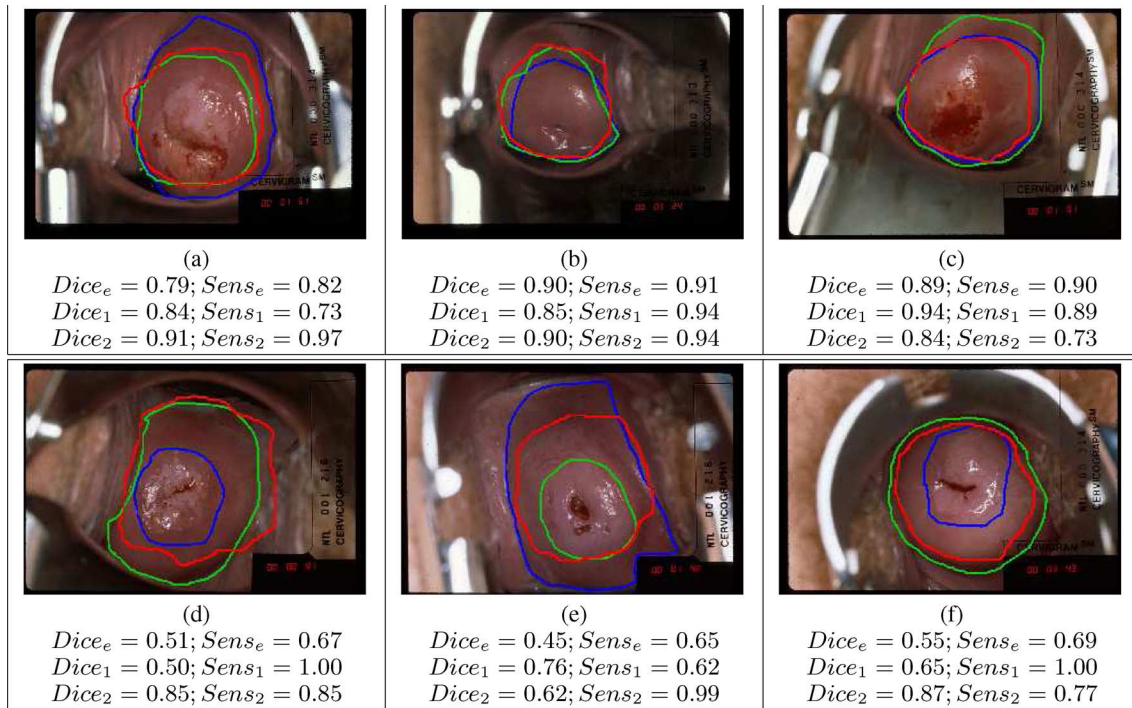|		
(a)	(b)	(c)
$Dice_e = 0.79; Sens_e = 0.82$	$Dice_e = 0.90; Sens_e = 0.91$	$Dice_e = 0.89; Sens_e = 0.90$
$Dice_1 = 0.84; Sens_1 = 0.73$	$Dice_1 = 0.85; Sens_1 = 0.94$	$Dice_1 = 0.94; Sens_1 = 0.89$
$Dice_2 = 0.91; Sens_2 = 0.97$	$Dice_2 = 0.90; Sens_2 = 0.94$	$Dice_2 = 0.84; Sens_2 = 0.73$
(d)	(e)	(f)
$Dice_e = 0.51; Sens_e = 0.67$	$Dice_e = 0.45; Sens_e = 0.65$	$Dice_e = 0.55; Sens_e = 0.69$
$Dice_1 = 0.50; Sens_1 = 1.00$	$Dice_1 = 0.76; Sens_1 = 0.62$	$Dice_1 = 0.65; Sens_1 = 1.00$
$Dice_2 = 0.85; Sens_2 = 0.85$	$Dice_2 = 0.62; Sens_2 = 0.99$	$Dice_2 = 0.87; Sens_2 = 0.77$

Fig. 14. Automated boundary detection (red) as compared to the markings of two experts (green and blue). Listed are the corresponding Dice and Sensitivity measures between the two experts ($Dice_e$, $Sens_e$) and between the algorithm and expert 1 and expert 2 ($Dice_1$, $Sens_1$, $Dice_2$, $Sens_2$, respectively).

TABLE III
QUANTITATIVE RESULTS FOR BOUNDARY DETECTION WITH AND WITHOUT A PRIOR SHAPE MODEL (Set_1, Set_2)

	Set_1		Set_2	
mean(std)	No Prior	Prior	No Prior	Prior
Dice	0.79(0.10)	0.81(0.10)	0.73(0.12)	0.75(0.13)
Sensitivity	0.95(0.05)	0.94(0.08)	0.96(0.07)	0.95(0.09)
FP	0.28(0.13)	0.13(0.12)	0.33(0.15)	0.16(0.14)
Hausdorff	254(105)	216(91)	300(125)	231(115)

TABLE IV
AVERAGE DICE AND SENSITIVITY MEASURES FOR A COMPARISON BETWEEN THE ALGORITHM AND THE RESULTS OF TWO EXPERTS—100 IMAGES (Set_3)

	Dice (mean)	Dice (std)	Sens (mean)	Sens (std)
expert1 vs expert2	0.88	0.1	0.9	0.07
algorithm vs experts	0.81	0.07	0.81	0.12

Dice and the Sensitivity measures, computed between the two experts and between the algorithm and each of the experts, are listed below corresponding images. The sensitivity between the experts was computed twice per image and averaged, using each of the experts as the ground truth.

Cervigrams (a)–(c) are examples of strong agreement between the expert markings. In these cases the automated cervix delineation results are consistent with these of the experts. In examples (d)–(f) there is a strong disagreement between the expert markings of the cervix region. The algorithm results are close to one of the experts, or within the range of their markings. Table IV summarizes the results for Set_3 images. Similar Sensitivity measures can be seen between the two experts and between the algorithm and the experts. This indicates that most of the cervix region, as marked by each of the experts, is detected by the algorithm. The average Dice result between the two experts is 0.88. The standard deviation of ± 0.1 indicates a strong variability between the experts. The average Dice measure between the algorithm and the experts is 0.81. Viewing the average Dice results between the experts as a desired objective, the automated algorithm achieves 92% of this objective.

B. Os Detection Results

Example results for os detection are presented in Fig. 15. In each image the os region is manually marked by the expert (white contour). The result of the automated os detection is marked by a green diamond.

Visual inspection indicates that in most cases the automatically detected os is located within the manually marked os region. Cases can be found where the os is misdetected, such as the right-most example in Fig. 15. In such cases, a prominent concave region that is located in the central part of the image does not necessarily correspond to the os. These misclassifications arise from a noncentered acquisition of the cervix region or from imprecise detection of the cervix boundaries.

A quantitative evaluation of the os detection process was performed using Set_1 and Set_2, for which a manual delineation of the os region by a medical expert is available. The minimal distance between the automatically detected os marker and the manually segmented os contour was found for each image in the selected set. A histogram of the distances obtained for the images in each of the sets is shown in Fig. 16. The distances are less than 10 pixels in 82% (98 out of 120) of the cases in Set_1 and for 79% (123 out of 158) of the cases in Set_2. The distance
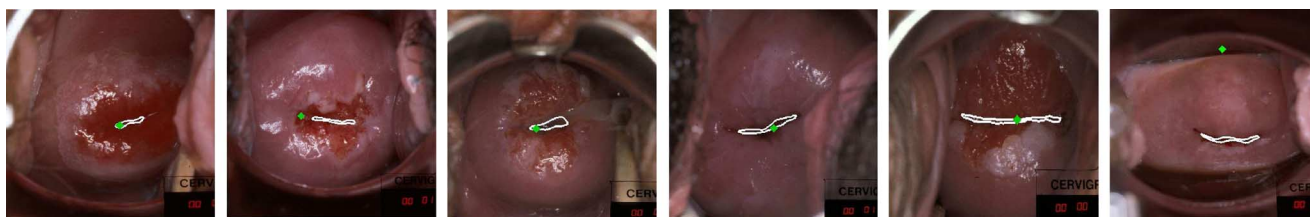
Fig. 15.   Os detection. Green diamond—automatic os detection. White—expert markings.
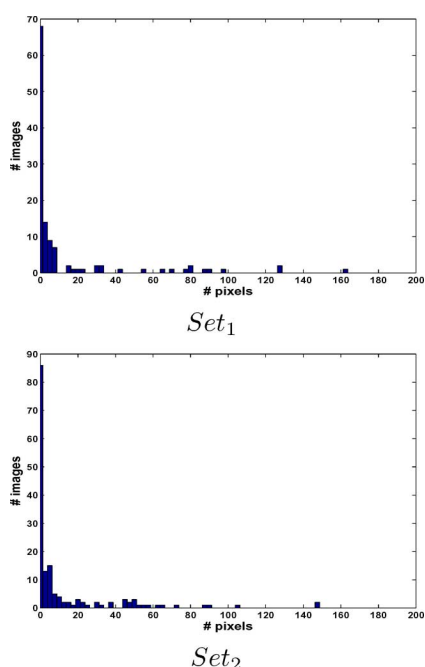

Set_1

Set_2

Fig. 16.   Histogram of the distances between the automatically detected os marker and the manually segmented os contour. x-axis: distance in pixels. y-axis: number of images per distance.

of 10 pixels is the approximate width of the os region, as marked by the expert.

C. Additional Cervigram Analysis Results

Fig. 17 presents an additional set of cervigram images, along with their automated analysis. Each row corresponds to a different cervix image. Column (a) shows the original cervigram with the coarse ROI marked. Column (b) shows the detection of SR, which is marked in black. Column (c) displays the same images following SR removal. Column (d) presents the refined cervix boundary extraction (white) and the expert marked boundary (blue). In column (e) results of os detection are given. The automatically detected os center point is marked in white, with the expert marking in blue.

The coarse ROI detection step (a) is very reliable. Large portions of irrelevant image regions are excluded while no information is lost. The detection of the specular reflection (b) was judged highly accurate in 90% of the cases in one test evaluation by a medical expert [17]; the subsequent filling step (c) produces a visually satisfactory, mostly specularity-free image. We note that in some cases small traces of specularities are left. Such a case can be seen in example (6) where a small bright

spot is present in the upper part of the cervix. In order to handle such cases, subsequent processing steps, such as the os detection, include postprocessing steps to discard any false positives that may be caused by undetected SR. The fact that even small residues of these bright spots cause misclassifications illustrates the importance of this preprocessing step.

The cervix boundary detection that is shown in (d) substantially improves the coarse ROI found in (a) by eliminating additional irrelevant regions that are close to the cervix. Results are very close to the human expert markings. The use of a shape prior imitates the human tendency towards marking a smooth, continuous, circular contour. Examples (2)–(4) show cases where the automatic contour tends to shrink inwards into the cervix region. This result is related to the prior shape of a circle which is currently used. In some cases, such as examples (5) and (6), the automatically detected boundary is larger than the boundary marked by the expert, indicating premature stopping of the evolution process as it evolves inwards from the coarse ROI boundaries. The os detection step (e) shows good results in a majority of these images.

VI. DISCUSSION AND CONCLUSION

This paper presents an automated analysis scheme for uterine cervix images, with an emphasis on the detection of two important landmarks: the cervix boundary and the os. The proposed scheme is hierarchical: it first defines a rough estimate for the cervix region within the image, and focuses the analysis within this region. In the second processing step, the specular reflection artifacts are detected and eliminated. A more exact delineation of the cervix boundary is then computed, along with detection of the os region. Each processing step is treated separately, using appropriate features and specific segmentation algorithms, with the goal of achieving robustness in the segmentation.

We introduce a novel approach based on geometrical curvature characteristics of the image for detection of the os and of the cervix boundary. This approach outperforms an existing clustering-based method in the task of cervix boundary detection. Incorporating prior shape information is shown to further improve the boundary detection quality.

In the proposed system we use an unsupervised approach for landmark detection. This unsupervised methodology focuses on each image as a unique entity, and analyzes and clusters features in the image, based on the specific (unique) content of that particular image. This approach is useful when a large variability exists across images in the given archive. In such cases, the cross-image variability imposes difficulties in learning a model per landmark type, or tissue type, that will be valid (generalizable) for the entire image archive.
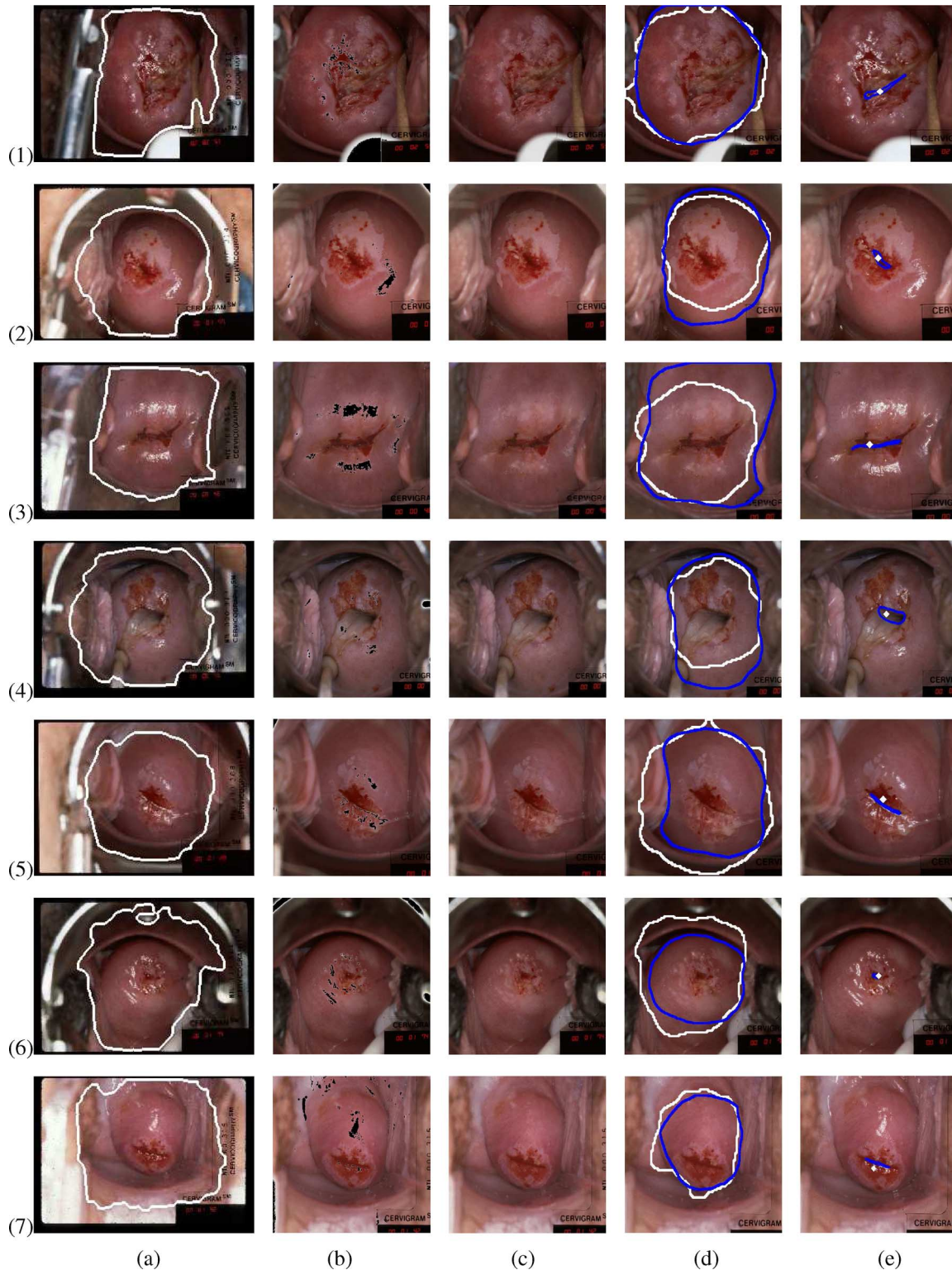
Fig. 17. Additional results. (a) Original cervigram with the initial ROI marked in white. (b) SR detection—black. (c) Cervigrams following SR filling. (d) Cervix boundary detection: Automated algorithm—white, manual contour—blue. (e) Os detection: Automatic algorithm—white, manual contour—blue.

Satisfactory results for the detection of the SR, the cervix boundary and the os, were achieved, thus constituting a good basis for further analysis of the cervigram content. Consistency in the results was shown across different image test sets. It should be noted that SR elimination is an important first step for a correct identification of the cervix boundary. Moreover,

detection of an accurate cervix boundary is critical for further landmark extraction. For example, the detection of the os landmark is dependent on the identification of its position relative to the cervix boundary.

The method used in this work for incorporating circular shape information, has the following advantages: 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 and in the case of a circular prior, computed only once. We are currently working on the incorporation of the more general, elliptical prior, into this framework. As the ellipse parameters (aspect ratio of the main axes) are not known in advance, its feature distribution has to be recomputed in each iteration. This increases the complexity of the process. Other methods are currently being developed in order to include the elliptical prior in a less complicated approach.

Future work will deal with the extension of the system and its validation in several directions: additional features, such as color and texture will be explored for the os detection task, and in the characterization of additional cervix tissue regions; within-image illumination correction and cross-image normalization schemes will be developed. Finally, we will incorporate the use of new sets of multiexpert data collected by NCI. Such data is critical for generating training sets, as well as for thorough computational validation of the system results.

REFERENCES

[1] J. Jeronimo, P. E. Castle, R. Herrero, R. D. Burk, and M. Schiffman, "HPV testing and visual inspection for cervical cancer screening in resource-poor regions," *Int. J. Gynecol. Obstetrics*, vol. 83, pp. 311–313, 2003.

[2] R. Herrero, A. Hildesheim, C. Bratti, M. E. Sherman, M. Hutchinson, J. Morales, I. Balmaceda, M. D. Greenberg, M. Alfaro, R. D. Burk, S. Wacholder, M. Plummer, and M. Schiffman, "Population-based study of human papillomavirus infection and cervical neoplasia in rural Costa Rica," *J. Nat. Cancer Inst.*, vol. 92, no. 6, pp. 464–474, 2000.

[3] R. Herrero, M. H. Schiffman, C. Bratti, A. Hildesheim, I. Balmaceda, M. E. Sherman, M. Greenberg, F. Cardenas, V. Gomez, K. Helgesen, J. Morales, M. Hutchinson, L. Mango, M. Alfaro, N. W. Potischman, S. Wacholder, C. Swanson, and L. A. Brinton, "Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project," *Pan American J. Public Health*, vol. 1, no. 15, pp. 362–375, 1997.

[4] M. Schiffman and P. Castle, "Human papillomavirus," *Arch. Pathological Lab. Medicine*, vol. 127, pp. 930–934, 2003.

[5] L. S. Massad, "American society for colposcopy and cervical pathology and the national institutes of health explore research collaboration," *J. Low Genit Tract Dis.*, vol. 10, no. 1, pp. 1–2, 2006.

[6] L. R. Long, S. Antani, J. Jeronimo, M. Schiffman, M. Bopf, L. Neve, C. Cornwell, S. R. Budhihas, and G. R. Thoma, "Technology for medical education, research, and disease screening by exploitation of biomarker in a large collection of uterine cervix images," in *Proc. 19th IEEE Symp. Comuter-Based Med. Syst. (CBMS)*, 2006.

[7] P. M. Cristoforoni, D. Gerbaldo, A. Perino, R. Piccoli, F. J. Montz, and G. L. Captianio, "Computerized colposcopy: Results of a pilot study and analysis of its clinical relevance," *Obstet. Gynecol.*, vol. 85, pp. 1011–1016, 1995.

[8] Q. Ji, J. Engel, and E. Craine, "Texture analysis for classification of cervix lesions," *IEEE Trans. Med. Imag.*, vol. 19, no. 11, pp. 1144–1149, Nov. 2000.

[9] B. W. Pogue, M. A. Mycek, and D. Harper, "Image analysis for discrimination of cervical neoplasia," *J. Biomed. Optics*, vol. 5, no. 1, pp. 72–82, 2000.

[10] H. Lange, "Automatic detection of multi-level acetowhite regions in RGB color images of the uterine cervix," *Proc. SPIE Med. Imag.*, vol. 5747, pp. 1004–1017, 2005.

[11] V. V. Raad, Z. Xue, and H. Lange, "Lesion margin analysis for automated classification of cervical cancer lesions," *Proc. SPIE Med. Imag.*, vol. 6144, pp. 1647–1659, 2006.

[12] 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," *Proc. SPIE Med. Imag.*, vol. 6144, pp. 1549–1556, 2006.

[13] Y. Srinivasan, F. Gao, B. Tulpule, S. Yang, S. Mitra, and B. Nutter, "Segmentation and classification of cervix lesions by texture and pattern analysis," *Int. J. Intell. Syst. Technol. Appl.*, vol. 1, no. 3/4, pp. 234–246, 2006.

[14] Z. Xue, S. Antani, R. Long, and G. Thoma, "Comparative performance analysis of cervix ROI extraction and specular reflection removal algorithms for uterine cervix image analysis," in *Proc. SPIE Med. Imag.*, 2007, vol. 6512, no. 65124I, pp. 1–9.

[15] Summary of safety and effectiveness data, optical detection system, LUMA (TM) cervical imaging system U.S. Food and Drug Administration, PMA P040028, 2006.

[16] G. Zimmerman, S. Gordon, and H. Greenspan, "Automatic landmark detection in uterine cervix images for indexing in a content-retrieval system," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2006, pp. 1348–1351.

[17] G. Zimmerman and H. Greenspan, "Automatic detection of specular reflections in uterine cervix images," in *Proc. SPIE Med. Imag.*, 2006, vol. 6144, pp. 2037–2045.

[18] C. M. Bishop, *Neural Networks for Pattern Recognition*. New York: Oxford Univ. Press, 1995.

[19] T. M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *Optical Soc. Amer.*, vol. 18, no. 11, pp. 2679–2691, 2001.

[20] R. C. Gonzalez and R. E. Woods, *Digital Image Process.*. Upper Saddle Rive, NJ: Prantice-Hall, 2002.

[21] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[22] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on the hamilton-jacobi formulation," *J. Computational Phys.*, vol. 79, pp. 12–49, 1998.

[23] L. He, Z. Peng, B. Everding, X. Wang, C. Y. Han, K. L. Weiss, and W. G. Wee, "Review: A comparative study of deformable contour methods on medical image segmentation," *Image Vision Comput.*, vol. 26, no. 2, pp. 141–163, 2008.

[24] R. Kimmel, "Fast edge integration," in *Geometric Level-Set Methods in Imaging, Vision and Graphics*. New York: Springer-Verlag, 2003.

[25] M. Farber, *Introduction to Differential Geometry* (in Hebrew). : Mea, 1999.

[26] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, *Statistical Shape Influence Geodesic Active Contours*, vol. 1, pp. 316–323, 2000.

[27] M. Rousson and N. Paragios, "Shape priors for level set representations," in *Proc. Eur. Conf. Comput. Vis.*, 2002, pp. 78–92.

[28] A. Tsai, A. J. Yezzi, W. M. W. , III, C. Tempany, D. Tucker, A. Fan, W. E. L. Grimson, and A. S. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, 2003.

[29] A. Litvin and W. C. Karl, "Coupled shape distribution-based segmentation of multiple objects," in *Proc. Inf. Process. Med. Imag. (IPMI)*, 2005, pp. 345–356.

[30] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distribution," *ACM Trans. Graphics*, vol. 21, no. 4, pp. 807–832, 2002.

[31] A. Tankus and H. Yeshurun, "Convexity-based visual camuflage breaking," *Comput. Vis. Image Understand.*, vol. 82, pp. 208–237, 2001.

[32] J. Jeronimo, L. R. Long, L. Neve, M. Bopf, S. Antani, and M. Schiffman, "Digital tools for collecting data from cervigrams for research and training in colposcopy," *J. Lower Genital Tract Disease*, vol. 10, no. 1, pp. 16–25, 2006.

[33] G. Gerig, M. Jomier, and M. Chakos, "Valmet: A new validation tool for assessing and improving 3d object segmentation," in *Proc. Med. Image Comput. Computer-Assisted Intervent. (MICCAI01)*, 2001, vol. 2208, pp. 516–523.