

CONTENT-BASED INDEXING AND RETRIEVAL OF UTERINE CERVIX IMAGES

Gali Zimmermann, Shiri Gordon and Hayit Greenspan

Department of Biomedical Engineering, Tel-Aviv University,
Tel-Aviv 69978, Israel

ABSTRACT

This work is motivated by the need for visual information management in the growing field of digital libraries and by the increasing information retrieval demands in the domains of medical imaging and telemedicine. We focus on a large database of digitized 35mm slides of the uterine cervix collected by the National Cancer Institute (NCI), National Institutes of Health (NIH), to study the evolution of lesions related to cervical cancer. As a first step towards this goal we focus on the problem of intelligently labeling (segmenting) regions of medical interest within the cervigram image. In this paper we use statistical tools for the segmentation of three tissue types of interest.

1. INTRODUCTION

Cervical cancer is the second most common cancer affecting women worldwide and the most common in developing countries. Although treatable and diagnosable, the available treatment can be complex and expensive in resource-poor regions. An alternative method of cervical cancer screening, termed cervicography, uses visual testing based on the color change of cervix tissues when exposed to acetic acid. This inexpensive method helps to detect abnormal cells that turn white (Aceto-White) after the application of the acetic acid [1]. The National Cancer Institute (NCI) has collected a vast amount of biomedical information related to the occurrence and evolution of uterine cervical cancer in a longitudinal multi-year study carried out in Guanacaste, Costa Rica. The data collected includes among other medical details, 60,000 cervicographic images ("Cervigrams"), as well as medical classifications for the Cervigrams into diagnostic categories. A major long-term objective is to develop a unique Web-based database of the digitized cervix images that will provide data for cervical cancer investigation.

This work is part of an on-going effort towards the creation of a content-based image retrieval (CBIR) system for the cervicographic images. Visual-based (i.e. content-based) indexing and retrieval based on the information contained in medical images is a newly emerging field that is expected to have a great impact on medical image databases. The challenge for vision researchers and engineers is to develop

tools for analyzing the content of images and to represent it in a way that can be efficiently searched and compared. The unique characteristics of medical images hinder the direct adaptation of content-based retrieval approaches, which are already in use for unstructured collections of images. A few systems have been recently introduced for medical image retrieval. These include text-based systems [2] and content-based systems like: the Automatic Search and Selection Engine with Retrieval Tools (ASSERT) and the Image Management Environment (IME) for high resolution CTs of the lung [3],[4].

In our research we seek to index the cervigram images with descriptors (parameters) derived from the image data in an automated fashion. We look for specific biomedical descriptors of high interest to the NCI research. In particular we focus on the following tissue types: The original squamous epithelium (SE), which is a featureless, smooth, pink epithelium; The columnar epithelium (CE) that extends between the endometrium and the squamous epithelium, and appears red and irregular; and the acetowhite (AW) region which is a transient, white-appearing epithelium following the application of acetic acid. Areas of acetowhitening correlate with higher nuclear density and are of clinical significance. A set of parameters can be extracted for each tissue; for example, for the AW tissue parameters can include the total area of the AW lesions in the image relative to the total image area, as well as a measure or measures of the color characteristics of the AW lesions. Such parameters would allow a user to retrieve images by size and color of the AW lesions.

In this paper we describe initial efforts for cervigram segmentation from which the regions of interest related with the desired tissues can be extracted, indexed and used by the CBIR system. The desired segmentation results are exemplified in Figure 1. A small number of works can be found in the domain of cervix image segmentation. Most of the works require the user to define regions of interest (ROI) on various cervix tissues [5], [6], [7]. Features such as color [6] and texture [7] are then extracted automatically. Manual lesion segmentation enables the extraction of geometric features as well [5]. Based on these features the different regions are classified to different cervix tissue types using var-

ious classifiers, such as neural networks [5] or the minimum distance classifier [7]. In a more recent work segmentation is performed based on the local frequency content within the image [8].

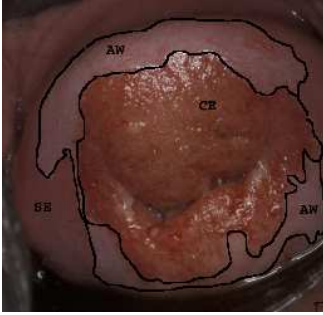


Fig. 1. Example for regions of interest in a cervigram: SE, CE and AW.

2. CERVIGRAM SEGMENTATION

We suggest a multi-stage segmentation process: Image pre-processing is first utilized to eliminate specular reflections. These reflections tend to interfere with the tissue segmentation task (section 2.1). Next, textured regions that are mostly related with the CE tissue are separated from the smooth regions of the AW and SE tissues based on a texture contrast feature (section 2.2). The smooth regions are then modeled as a mixture of Gaussians in a color feature space using a recently proposed continuous image representation [9]. Following the model generation, probabilistic image segmentation is enabled (section 2.3).

2.1. Preprocessing - Identification of specular reflections

Specular Reflections (SR), or highlights, are generated in every cervigram as a consequence of the image acquisition process (see Figures 1 and 2). They appear as small and very bright saturated regions in various parts of the image. Modeling of the SR regions is a difficult task. Pixels of SR are identified as other tissue types, such as AW regions. The indexing process is thus hindered. For example, an image that doesn't contain AW regions may be incorrectly labelled as one which contains AW. In order to overcome this problem, the specular reflections are removed in a pre-processing step, before the tissue modeling process.

Pixels with SR have very high intensity (V) and low saturation (S) values [10]. Regions of SR appear as small bright spots, and usually cause large gradients around them. We locate regions with high gradients, high intensity and low saturation as our regions of interest in the SR detection phase. We model the distribution of the pixels within these regions as a mixture of Gaussians in an $S - V$ feature space

and classify them into SR and non-SR pixels using a probabilistic segmentation procedure, as described in section 2.3.

Figure 2 shows results of SR detection. On the left is the original image. On the right the SR regions are masked out by the algorithm and marked as black regions.

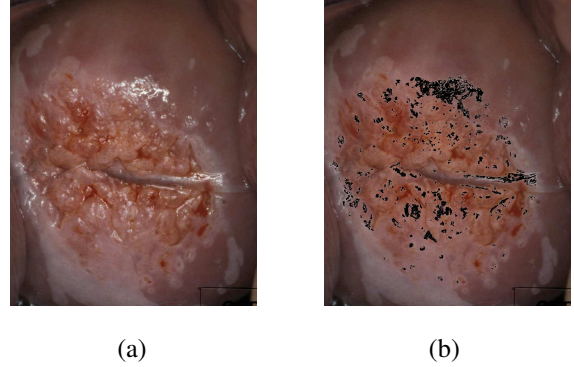


Fig. 2. Examples of specularities masking: (a) Original cervigram; (b) Masking out the specular reflections (marked in black).

2.2. Texture Segmentation

The columnar epithelium (CE) appears as a coarse and textured region, which contains a mixture of all the colors in the cervigram image. The rest of the image is relatively smooth and non-textured. Texture features are thus used for the segmentation of the CE region (using color features for this task would result with over-segmentation of the CE region).

The texture features used describe both the underlying texture parameters and the adequate texture scale [11]. The scale is defined as the width of the Gaussian window within which gradient vectors of the image are pooled. The second moment matrix for the vectors within this window, computed about each pixel in the image, can be approximated using Equation 1:

$$M_{\sigma}(x, y) = G_{\sigma}(x, y) * (\nabla I)(\nabla I)^T \quad (1)$$

where G_{σ} is a separable binomial approximation to a Gaussian smoothing kernel with variance σ^2 , and (∇I) is the gradient of the image intensity. Two texture descriptors are extracted for each pixel: polarity and texture contrast. Polarity is a measure of the extent to which the gradient vectors in a certain neighborhood all point in the same direction, as defined in Equation 2:

$$p_{\sigma} = \frac{|E_{+} - E_{-}|}{E_{+} + E_{-}} \quad (2)$$

where σ is the scale. The definitions of E_{+} and E_{-} are:

$$\begin{aligned}
E_+ &= \sum_{x,y} G_\sigma(x,y) [\nabla I \cdot \hat{n}]_+; \\
E_- &= \sum_{x,y} G_\sigma(x,y) [\nabla I \cdot \hat{n}]_-
\end{aligned} \quad (3)$$

where $[\cdot]_+$ and $[\cdot]_-$ are the rectified positive and negative parts of their argument and \hat{n} is a unit vector perpendicular to ϕ (the direction of principal eigenvector of the second moment matrix, as defined in Equation 1). This feature is used later for the selection of an appropriate texture scale for each pixel in the image. The contrast relates to the energy of the gradients in the vicinity of each pixel as given by Equation 4, where λ_1 and λ_2 are eigenvalues of M_σ ($\lambda_1 \geq \lambda_2$) at each location:

$$contrast = 2\sqrt{\lambda_1 + \lambda_2} \quad (4)$$

The process of selecting an appropriate scale is based on the derivative of the polarity with respect to the scale. For each pixel (x, y) the scale is selected as the first value for which the difference between values of polarity at successive scales is less than 2%.

At the end of the texture features extraction phase, each pixel is associated with a contrast feature of the appropriate scale. Using a hard threshold on the contrast feature value we separate between textured and non-textured pixels, thus segmenting the image into textured and non-textured regions. The textured regions are then masked out and labelled as CE regions.

2.3. Statistical modeling and segmentation of smooth regions

Statistical modeling and segmentation, based on color features, is performed on the non-textured regions of the cervigram. During the segmentation process each pixel within the smooth regions is represented with a three-dimensional color descriptor in the *Lab* color space which was shown to be approximately perceptually uniform [9]. Pixels are grouped into homogeneous regions by grouping the feature vectors in the selected three-dimensional feature space. The underlying assumption is that the smooth region colors are generated by a mixture of Gaussians. Each homogeneous region in the image plane is thus represented by a Gaussian distribution, and the set of regions in the image is represented by a Gaussian mixture model.

The distribution of a d -dimensional random variable is a mixture of k Gaussians if its density function is :

$$f(y) = \sum_{j=1}^k \alpha_j \frac{1}{\sqrt{(2\pi)^d |\Sigma_j|}} \exp\left\{-\frac{1}{2}(x - \mu_j)^T \Sigma_j^{-1} (x - \mu_j)\right\}, \quad (5)$$

where d is the feature space dimension, α_j are the probabilities of the occurrence of each Gaussian, and μ_j, Σ_j are the mean and the covariance matrix of each Gaussian cluster respectively. The Expectation-Maximization (EM) algorithm [9], is used to determine the maximum likelihood parameters of a mixture of Gaussians in the selected feature space. In our case, $d = 3$, and the learned mixture is composed of two Gaussians, one representing the SE tissue and the other representing the AW tissue.

An immediate transition is possible between GMM representation and probabilistic tissue segmentation. Each pixel of the original image is represented by the three-dimensional color feature vector x . The labelling of a pixel related to the feature vector x is chosen as the maximum *a posteriori* probability:

$$label(x) = \arg \max_j \alpha_j f(x | \mu_j, \Sigma_j) \quad j = 2 \quad (6)$$

i.e. each pixel gets affiliated with the most probable Gaussian cluster in the learned GMM. Following this segmentation step pixels belonging to the smooth (non-textured) regions are classified into AW and SE regions as desired.

3. RESULTS

Unsupervised image segmentation using the described procedure was performed on cervigram images. Figure 3 shows a segmentation example. The original image is shown in Figure 3(a). Segmentation results are shown in Figure 3(b). The colors used in the segmentation map are synthetic. From lighter to darker colors they represent the AW, SE and CE tissues. Figure 3(c) displays the AW regions marked on the original image. In this result the specularities were replaced with the average color of their neighborhood. A good separation between the three tissues of interest can be seen. Most of the AW regions were detected correctly. Some small regions on the outer parts of the image were misclassified as CE regions. These regions can be correctly re-classified in a post-processing step, where we can incorporate high level knowledge on the cervigram layout.

4. DISCUSSION

In this work we present initial segmentation results for cervigrams using GMM modeling and probabilistic image segmentation. In particular we focus on three tissue types: the squamous epithelium (SE), the columnar epithelium (CE) and the acetowhite lesions (AW). The AW regions are of particular interest since estimates of their size are of clinical significance. The results, although preliminary, look promising for the differentiation of the three examined tissue types. We are currently working on combining the seg-

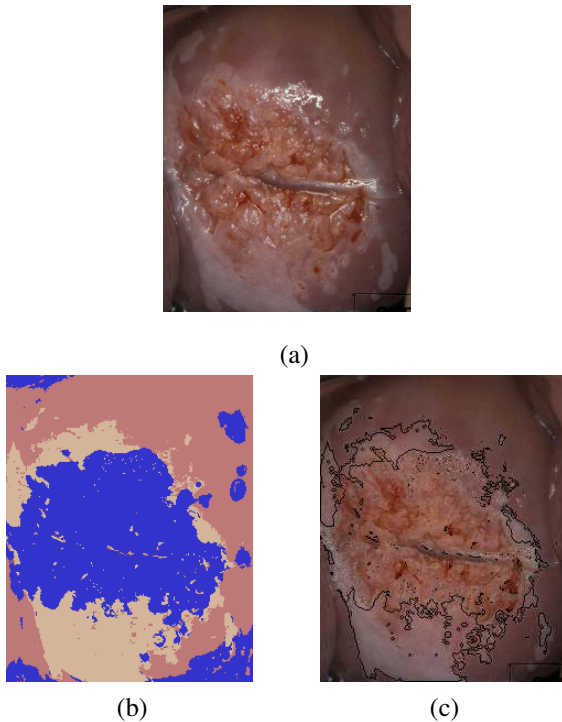


Fig. 3. Segmentation example: (a) Original image; (b) Segmentation map; (c) Aceto-White regions marked on the original image.

mentation steps of removing specular reflections, and modeling in color and texture space, using advanced segmentation schemes. Future work entails the incorporation of additional features to the tissue representation. The evaluation of the segmentation results by medical experts is still required. A long term task includes using the segmentation results for indexing of cervigrams, and AW regions in particular, as a part of a complete cervigram CBIR system.

Acknowledgement

We would like to thank the Communications Engineering Branch, National Library of Medicine, NIH, for the data and support of the work.

5. 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," *International Journal of Gynecology and Obstetrics*, vol. 83, pp. 311–313, 2003.
- [2] L.R. Long, S.R. Pillemer, R.C. Lawrence, G.H. Goh, L. Neve and G.R. Thoma, "Webmirs- web-based medical information retrieval system," in *Proc. of SPIE*, vol. 3312, pp. 392–403, 1998.
- [3] C.R. Shyu, C.E. Brodley, A.C. Kak, A. Koska, A.M. Aisen and L.S. Broderick, "Assert - a physician-in-the-loop content-based retrieval system for hrct image database," *Computer Vision and Image Understanding*, vol. 75, pp. 111–132, 1999.
- [4] A.F. Abate, M. Nappi, G. Tortora and M. Tucci, "Ime - an image management environment with content-based access," *Image and Vision Computing*, vol. 17, pp. 967–980, 1999.
- [5] P.M. Cristoforoni, D. Gerbaldo, A. Perino, R. Piccoli, F.J. Montz and G.L. Captiano, "Computerized colposcopy: Results of a pilot study and analysis of its clinical relevance," *Obstet. Gynecol.*, vol. 85, pp. 1011–1016, 1995.
- [6] B.W. Pogue, M.A. Mycek and D. Harper, "Image analysis for discrimination of cervical neoplasia," *Journal of Biomedical Optics*, vol. 5, no. 1, pp. 72–82, 2000.
- [7] Q. Ji, J. Engel and E. Craine, "Texture analysis for classification of cervix lesions," *IEEE Transaction on Medical Imaging*, vol. 19, no. 11, pp. 1144–1149, 2000.
- [8] V. Van Raad, "Frequency space analysis of cervical images using short time fourier transform," in *Proceedings of the IASTED International Conference of Biomedical Engineering*, vol. 1, no. 1, pp. 77–81, Jun 2003.
- [9] H. Greenspan, J. Goldberger and L. Ridel, "A continuous probabilistic framework for image matching," *Journal of Computer Vision and Image Understanding*, vol. 84, pp. 384–406, 2001.
- [10] T.M. Lehmann and C. Palm, "Color line search for illuminant estimation in real-world scenes," *Optical Society of America*, vol. 18, no. 11, pp. 2679–2691, 2001.
- [11] C. Carson, S. Belongie, H. Greenspan and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, 2001.