Image Segmentation of Uterine Cervix Images

Antonios Perperidis

Cervical cancer is one of the most common forms of cancer afflicting women [1, 2]. Cervical cancer can be cured if it is detected during its early stages and treated appropriately [3]. The traditional method of detecting cervical cancer is the Pap smear test [4]. However, the Pap smear test is only a screening test and not a diagnostic tool [5]. In addition, it is not very accurate technique that requires trained personnel to evaluate the results.

Digital Colposcopy is an inexpensive, non-invasive diagnostic tool [5]. In colposcopy, the cervix is examined by using a specially designed binocular microscope, which is called colposcope [5]. In colposcopy, the abnormal cervical regions appear to be whiter than the original cervix and are called acetowhite [5]. Modern colposcopes can produce a digital image of the cervix [5]. However, there is currently an absence of specialized image processing software that has the ability to process automatically the images acquired in colposcopes [5]. Therefore, trained personnel is still required for the evaluation of results.

Recently, there have been many attempts to study the colposcopic image, extract features and automatically detect abnormalities. This document provides a description of a technique [6] that focuses on color and texture features and attempts to segment the cervical image in the following tissue types [6]:

- The original squamous epithelium (SE): Featureless, smooth, pink epithelium.
- The columnar epithelium (CE): Red, irregular.
- The acetowhite (AW): Transient, white-appearing epithelium. (The size of this region is of clinical significance)

The cervical segmentation technique proposed by Gordon et al. [6] consists of two stages, the feature selection and extraction stage and the statistical modeling and segmentation stage.

Feature Selection and Extraction

Images are modeled as a mixture of Gaussians in a color-texture feature space. Each pixel is represented by a 3-D color descriptor in the Lab color space. This approach extracts color features from the cervical image. Color features are very useful in the analysis of cervical images. Fig 1(b) demonstrates the apparent separation in the color space between the pixels of the three tissues of interest that are selected in Fig 1(a). The clusters within the tissue color distribution can be modeled with Gaussian or Gaussian mixture models. However, as we can see from Figure 1(b), color does not provide all the required information for the analysis of the cervical images. Additional features are required in order to be able to distinguish between the tissues of the cervical.

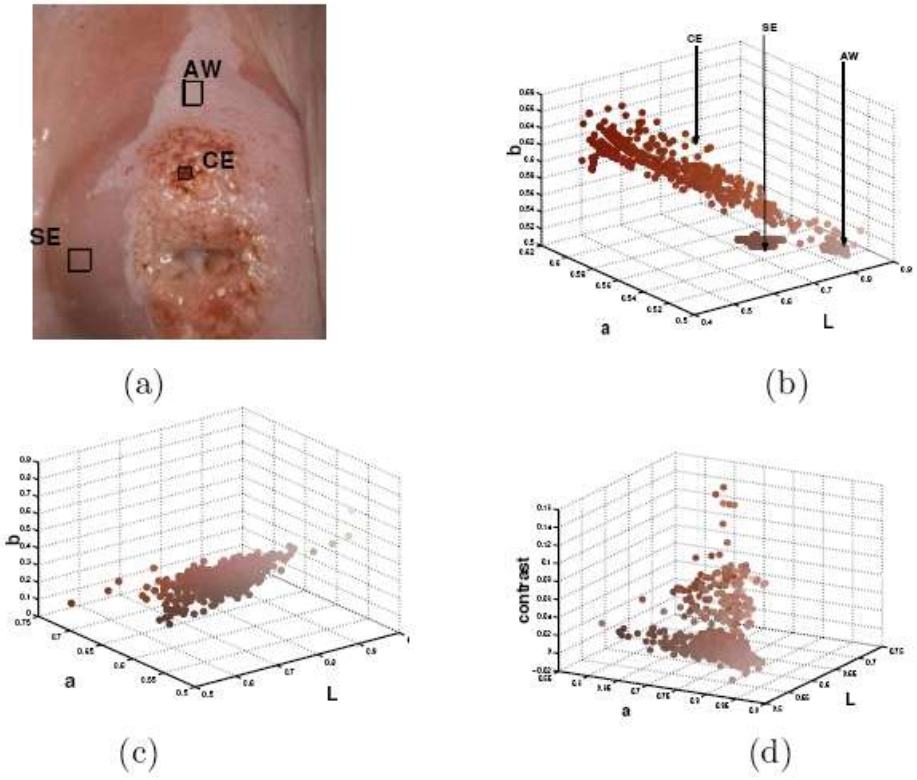

Figure 1 [6]. Pixels scatter in the feature-space. (a) Cervigram image, with manually selected patches from each tissue; (b) Pixels from each patch in Lab color-space; (c) Pixels from entire image in Lab color-space; (d) Pixels from entire image in La-contrast color-texture space.

Texture features are used to describe the texture parameters and the texture scale of the cervical image [6]. Two texture descriptors are extracted for each pixel:

- The polarity: a measure of the extent to which the gradient vectors in a certain neighborhood all point in the same direction [6].
- The texture contrast: related to the energy of the gradients in the vicinity of each pixel [6].

The derivative of the polarity with respect to the scale is used in oder to select an appropriate scale [6]. Then, the contrast feature of the appropriate scale is extracted and used during the segmentation process [6]. A separation between textured and non-textured clusters can be seen on the contrast axis of the combined color-texture space (Figure 1(d)).

Statistical Modeling and Classification

After the feature extraction stage, each pixel is represented with a four-dimensional feature vector [6]. This feature vector is used for grouping pixels into homogeneous regions.

Assuming that the image colors and texture are generated by a mixture of Gaussians then [6]:

- Each homogeneous region in the image plane is represented by a Gaussian distribution.

- The set of regions in the image is represented by a Gaussian mixture model.

The three Gaussian clusters that correspond to the three tissue types of interest are determined by using the Expectation-Maximization (EM) algorithm [6]. Then each pixel of the original image can be related with most probable Gaussian cluster by using the maximum a posteriori probability [6]:

$$label(x) = \arg\max_{j} \ \alpha_j f(x|\ \mu_j, \Sigma_j)$$

where 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.

Results

The results of the described approach look promising for the differentiation of the three examined tissue types of interest we described earlier in this document. Improvements on the the initial results include:

- the investigation on additional features [6].
- the elimination of specular reflections [6].
- the merging of different regions related to the same tissue [6].

References

[1] American Cancer Institute, Cancer Facts and Figures 2005.

[2] http://cancer-symptoms.org, Cervical Cancer Information.

[3] http://www.nci.nih.gov/, News Article - Treatment Options for Cervical Cancer by Stage.

[4] National Cancer Institute, Cancer Facts - The Pap Test: Questions and Answers, http://cis.nci.nih.gov/fact/5_16.htm

[5] Women's Health at www.about.com, News article - What is Colposcopy?, http://womenshealth.about.com/cs/cevicalconditions/a/colposcopy.htm.

[6] Shiri Gordon, Gali Zimmerman, Hayit Greenspan, "Image Segmentation of Uterine Cervix Images for Indexing in PACS," 17th Symposium on Computer-Based Medical Systems, Bethesda, Maryland, June,2004.