Lesion Margin Analysis for Automated Classification of Cervical Cancer Lesions

Viara Van Raad*, Zhiyun Xue and Holger Lange
STI – Medical Systems, 733 Bishop St, Makai Tower, Suite 3100, Honolulu, HI, USA 96813
Telephone: +1 808 5404771, Fax: +1 808 5404850, e-mail: vraad@sti-hawaii.com

ABSTRACT

Digital colposcopy is an emerging technology, replacing the traditional colposcope for diagnosis of cervical lesions. Incorporating automated algorithms within a digital colposcopy system can improve the reliability and the diagnostic accuracy of cervical precancer and cancer. An automated computer-aided diagnosis (CAD) system can assess the three important cervical diagnostic cues: the color, the vascular patterns and the lesion margins with quantitative measures, similar to the way colposcopists use the Reid's index in traditional colposcopy. In this work we present a novel way to analyze and classify the global and the local features of one of the three major components in colposcopy diagnosis – the lesion margins. The margins of cervical lesion can be described as 'feathered,' 'geographic,' 'satellite,' 'regular or smooth' and 'margin-in-margin,' or they can be of mixed type. As margin characterization is a complex task, we use irregularity descriptors such as compactness indices and curvature descriptors. To address the complexity of the problem, the dependency of scale and the position of the lesion on the cervical image, our method use novel Fourier energy descriptors. The conceptually complex analysis of describing lesions as 'satellite' lesions or lesions with multiple margins is performed using descriptors, where the distance, the position and the local statistical estimates of image intensity play important role. We trained this new algorithm to classify and diagnose the cervix, evaluating only the lesions. The accuracy of the results is assessed against a 'ground truth' scheme, using colposcopists' annotations and pathology results. We report the resulted accuracy of the classification method assessed against this scheme.

Keywords:
Cervical cancer detection in vivo; Computer-Aided Diagnosis in colposcopy; signal processing, geometric model-based and statistical methods;

1. INTRODUCTION

1.1. AUTOMATED DETECTION OF CERVICAL PRECANCER AND CANCER

Each year, 371,000 new cases of invasive cervical cancer are diagnosed worldwide[1]. Cervical cancer is the most frequent type of cancer in women under 35 years of age[2]. It is a preventable disease, one in which the neoplastic changes grow slowly and they can often be treated successfully in the early stages of neoplasia. Visual impressions of the cervix are highly correlated with the accuracy of staging and diagnosis of cancer[3], therefore a computer aided diagnosis and digital image analysis is a promising evaluation tool for the cervix *in vivo* [4-5]. Digital imaging technology allows capturing images during visual examination of the cervix and after the application of 5% of acetic acid on the cervix – a technique introduced by Dr. Hans Hilsenman in 1925[10]. Acetic acid triggers a temporary whitening of the tissues with an 'off white' color. This is often called acetowhitening or acetowhite (AW) and the phenomena is due to a progressive and reversible alteration of the light scattering properties of the abnormal epithelium as a result of the existing compositional and/or functional alterations[6]. The AW appears on the tissues that are in the process of metaplasia (a healthy transformation of one tissue into another) or dysplasia (unhealthy change of a tissue). Some of the AW areas that remain longer indicate higher grade of dysplasia. These areas also have distinctive strong white colors and they are also known as AW lesions. AW is an important indicator of cervical cancer precursor. The attributes of the AW lesions can be assessed further in order to stage the precancer. Colposcopists use an indexation for each one of the cancer indicators and this scoring system is known as Reid's index [7]. Reid's colposcopic index is one of the earliest and most effective scoring systems for evaluating cervical cancer and precancer. Using Reid's scoring systems colposcopists attempt to quantify the cervical precursors and to provide an objective assessment of the cervix.

*Corresponding author

A digital colposcope is used to capture cervical images during the colposcopic examination while colposcopists assess the cancer precursors. Texture, AW lesions' boundaries and vascular patterns are very important cues in the visual evaluation of cervical precancer and cancer. The most important cancer precursors assessed in Reid's scoring systems are the AW lesions, their margins and the superficial vessel patterns. In Reid's scoring system each one of these acquire 0-3 points. The stage of colposcopically diagnosed disease is lower if the Reid's score is low[7].

In this paper we study the AW lesions and their margins in terms of their piecewise smoothness and their spatial relationship to one of the most important areas of the cervix – the Transformation Zone (TZ) (as this is the area where the cervical cancer occurs frequently). TZ surrounds the centrally located Cervical Canal (CC). TZ also contains the 'transforming columnar' (aka 'metaplastic') into Squamous Epithelium (SE) (Figure 1).

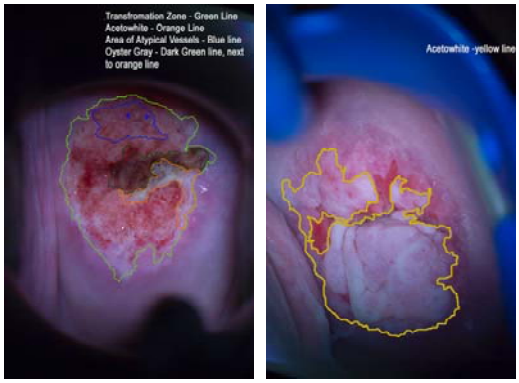

Figure 1. Left images shows a delineation of TZ – light green line; AW with internal irregular margins and smooth external margins – orange line; area of atypical vessels – blue line and 'gray' colored AW – dark green delineation; Right image shows a large AW lesion with distinct Lesion Margins (LMs) – the left portion has a smooth outline, whilst the right portion has a highly irregular outline.

Cervical images with high grade dysplasia are shown in Figure 1 (left and right). The left image has an AW lesion encircled by orange delineation; the Lesion Margins (LMs) of the AW lesion and the corresponding atypical vessels populated with mosaics are prominent within the TZ in this image. The right image on Figure 1 shows a prominent acetowhitening as a 'flat white' lesion. The smooth portion of the AW lesion is located on the left hand side of the lesion and the irregular portion of the LM is the right side of the lesion. Figure 2 depicts an original (left) and annotated (right) cervical image with transparent AW lesions, located centrally.

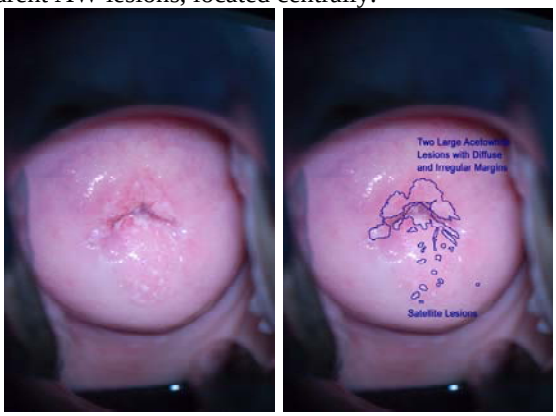

Figure 2. An original cervical image with low grade dysplasia (left). Right image shows delineated AW lesion with diffusive satellite portions. The AW has highly irregular and diffusive LMs, pointing to low grade dysplasia.

The cervical images from each one of 100 patients were captured during a colposcopic examination during the clinical trials conducted in Lima, Peru, where women with positive cytology test results or having symptoms of cervical disease

were referred for colposcopic examination. The images were taken under controlled illumination conditions and controlled timing (5 seconds apart) after acetic acid (5%) application. Later these images were subjected to image processing to extract meaningful features for automated computerized diagnostic and support.

1.2. ACETOWHITE LESION MARGINS AND THEIR SIGNIFICANCE

The AW lesions and their associated features, such as color, texture and borders can be evaluated and assessed quantitatively using signal and image processing methods [4]. Special attention is paid to the LM irregularity. An automated characterization of the lesions' borders is one of the major goals of the work described here. Another goal is the automated detection of the AW lesion. Therefore, the understanding of the lesion margins and their relation to the staging of cervical cancer is a key point of such analysis and characterization.

Lesion margins of the AW lesions on the cervix have a variety of forms and many visual interpretations. In order to categorize the AW lesions, the classification of LM can be grouped into two major categories, each one of those categories associated with *low* and *high* grades of cervical intraepithelial neoplasia (CIN). Low grade is known as mild dysplasia (CIN I), while high grade includes moderate dysplasia (CIN II), severe dysplasia (CIN III – cancer *in situ*), and cancer.

Lower indexation by Reid's index (1-2 points) indicates low grade lesions (CIN I). Reid indices are cumulative for color, margins and vessel pattern. Reid's indexation yields a low score for 'indistinct borders,' known as diffuse borders, irregular (e.g. 'angular lesions' or 'jagged'), or 'satellite lesions' [7]. Higher indexation (3-5 points) indicates high grade lesions. Reid's indexation for 'regular' ('smooth') lesions with a 'straight' outline is an indication for higher grade dysplasia and cancer (CIN II and CIN III). Strong demarcation ('distinct') LM is a well-defined characteristic of CIN III or sometimes as cancer *in situ*, while the diffusivity of the LMs indicates low grade lesions.

The perceptual and qualitative indications of 'irregular,' 'jagged,' 'satellite,' 'diffuse' and 'distinct' margins are hard to quantify, so a knowledge-based system that describes these terms must be introduced. First, training data is needed for the analysis of the lesion margins of the AW so an automated system of LM characterization can be developed. We approached the characterization of lesion margins aiming to achieve staging of a cervical lesion and automatically diagnosing the patient by the information from the LMs characterization only. In the current study we aim to determine the piecewise *irregularity* of the contour, which points to the *low stage* of precancer. If a *smooth* segment of the contour is present, this points to *high grade dysplasia*. If the contour contains only irregular segments, the detected AW lesion can be classified to be possibly a low grade. Using the quantified information for the LMs of AW an automated support/decision in clinical diagnosis of cancer during colposcopy examination will be offered.

Previously, Claude[8] and her co-workers performed neural network classification using features extracted from colposcopy images acquired after application of Lugol solution on the cervix. In certain cases colposcopists use Lugol's solution to define the AW[7] but acetic acid application is mandatory during colposcopic exam. This image analysis work involved a characterization of 'sharp contours', 'blurred contours with small regions,' 'blurred contour with fuzzy edges' and 'doubtful contour.' The AW contours obtained after Lugol's solution were assessed by examining the topography of the contour lines cut across the image contour with lines positioned in radial direction. Contours were analyzed in terms of their topology changes such as contours' peaks. The training features were extracted and used in the training phase of two networks: a Multilayer Perceptron (MLP) and a Probabilistic Network based on Parzen estimation (PNN). A k-Nearest Neighbors (kNN) algorithm was designed as well to be used as a 'reference method' for comparison of the performance of the MLP and PNN. Both MLP and PNN yielded similar results and a successful classification was achieved. The work presented by Claude et al.[8] is one of the pioneering works for integrating the elements of Reid's index into automated diagnostic support for the challenging problem of classification of cervical precancer and cancer[9]. Nevertheless that Claude et al.[8] method was successful, in the current work we aim to classify the contours of the AWs obtained as a result of the application of acetic acid, because the acetic acid application is the most common practice in colposcopy and Lugol's application is an optional.

The balance of this paper is organized as follows. Section 2 contains a description of the statistical method known as Maximum A- Posteriori (MAP) optimization method for segmentation of AW lesions. Section 3 contains a detailed explanation of the method for partition of the closed contour, obtained after the segmentation of the AW lesions,

followed by description of the characterization methods of each one of the segments. Section 4 contains the description of the data and data's preprocessing steps. Section 5 reports the results of the AW segmentation and piecewise curve characterization. Section 6 contains discussion points and conclusions.

2. DETECTION OF AW LESIONS USING MAXIMUM –A –POSTERIORI (MAP) ALGORITHM

Although a very strong sign of dysplasia, AW lesions have many subtle features, such as color saturation in terms of opacity, transparency, texture, raised borders and diffusivity of the AW. There are many subjective evaluations with strong perceptual values along AW lesion characterization. All of them are based on how Human Visual System (HVS) learns and processes visual information. Therefore, given that not all types of acetowhitenings are lesions, an automated detection of AW lesions is a challenging task. The AW detection in this paper is for the 'strong white' color, often associated with high grade dysplasia.

In the current experiments the automated detection of the AW lesions and lesion borders is accomplished by the application of a statistical method that belongs to the Bayesian classifier methods, also known as Maximum – *a – Posteriori* feature detection (MAP) [10]. MAP method uses a simplified analysis modeled by a Gaussian Mixture Model (GMM), probability estimation achieving bimodal segmentation of the image. As a result of MAP's successful segmentation, the AW lesion area in the resultant binary image is marked with 'ones,' while the area of the rest of the image is marked with 'zeros.'

The aim is to detect the AW lesion, using an optimization algorithm (maximum-*a posteriori* algorithm or MAP) for bimodal segmentation in the RGB color-space, estimating the values of the AW as a pixel set. The implementation of the MAP algorithm is accomplished by using a classical Bayesian approach. The segmentation method is based on pre-defined estimates of the GMM values [10] for the AW lesions taken from the adjacent areas in images from a training set. The case of AW lesion detection is a case of finding an 'object' with an 'off white' color opposite to the rest of the tissues and paraphernalia in the image, considered as the 'background,' represented as a mixture of two classes as bi-variate Gaussian Mixture Model (GMM). The two classes are A_1 and A_2 - respectively for the AW – labeled as region S, and its complimentary region - S^C. A model is created by using a training set of images T.

The probability density function (PDF) of the Gaussian Mixture Model is to be described as:

$$P(\vec{x}|\omega_i) = \left(\begin{array}{l} \frac{1}{(2\pi)^K \left|\Sigma_{\omega_1}\right|^{1/2}} \cdot \\ \exp\left(-\frac{1}{2}(\vec{x} - \mu_i)^T \Sigma_{\omega_1}^{-1} (\vec{x} - \mu_i) \right) \end{array} \right) \quad (1)$$

Where the Σ_i is the collective covariance matrix of the pixels that belong to the three-dimensional training set of T images with mean vectors μ_i and the corresponding pixels $\vec{x}_i$ representing the S and S^c, respectively for $i = 1,2$. The term $M_i = \frac{1}{2}(\vec{x} - \mu_i)^T \Sigma_\omega^{-1} (\vec{x} - \mu_i)$ is recognized as the Mahalanobis distance for i. The distance between the two clusters is:

$$d_i(x) = \ln P(\omega_i) - 0.5 \ln 2\pi - 0.5 \ln\left|\Sigma_{\omega_i}\right| - \frac{1}{2}(\vec{x} - \mu_i)^T \Sigma_\omega^{-1} (\vec{x} - \mu_i) \quad (2).$$

Segmentation is achieved by examining the adjacent values, corresponding to each one of the pixels' in the image, tested if it is closer to the one of the two clusters A_1 or A_2, represented in three dimensional color space. Each one of the classes A_1 or A_2 form specific cluster (range of values) represented by Σ_i, μ_i. The values are determined after a collection of pixels from S and S^C, respectively (AW and 'background').

The measures are obtained by calculating the Mahalanobis distance for each pixel of the RGB image $f(\alpha,\beta)$. The probability of classifying the pixel for A_1 or A_2 will be increased if their probabilities of occurrence ρ_1 and $\rho_2 = 1-\rho_1$ are known. In logarithmic terms, the optimum statistical classifier will find optimal distance measure considering that $\rho_i = \ln P(\omega_i)$ and $\rho_1 + \rho_2 = 1$. The estimated distance is:

$$\Delta = 2\ln\left[\frac{\rho_1|\Sigma\omega|}{(1-\rho_1)|\Sigma\omega|^C}\right] \quad (3)$$

The pixel belongs to AW area ($x_i \in AW$) if the value of the pixel included in the calculation for the differences in the two Mahalanobis distances fulfils the inequality:

$|M_1 - M_2| < \Delta$ (4).

The inequality (4) is the classification condition. MAP algorithm's first part is the estimation of the parameters illustrated in the flow chart in Figure 3 and follows by the detection phase using the parameters from the estimation part fulfilling the inequality (4). The first part of the MAP (the estimation) is finding the probability of the occurrence of the pixels within the AW region and its complementary, the related covariance matrices and mean vectors.

The second part is the comparison of the Mahalanobis values for each pixel under a test with the calculated 'distance' threshold assessing the validity in the inequality in (4). A detailed description of the method for detection of cervical features using MAP algorithm is described by Van Raad[11].

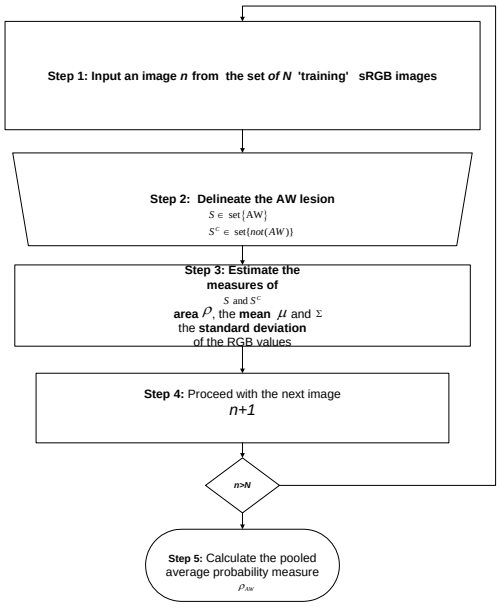

Figure 3. Algorithm flow for the estimation part of MAP algorithm for bimodal segmentation. The method involves a training set of images. Step 2 can be replaced by an overlying of binary mask on the color image for automated selection of pixel values.

3. CHARACTERIZATION OF THE LESION MARGINS

The lesion margin characterization can be viewed as a problem of analysis or characterization of a curve that encircles the AW lesion, previously segmented by MAP algorithm (result captured as a binary image). The first task is to segment the curve's component in a meaningful way; the second task is to characterize the segmented components. Therefore, for the development of the characterization method two questions are raised – the method for segmenting the curve into meaningful components, followed by characterization of those 'segments' as 'smooth' or 'irregular.'

In a case of multiple AW lesions, they are analyzed in terms of their mutual spatial relation and the area that each lesion occupies. The smoothness/irregularity of the LM contour is perceptually easy to understand and explain, while it seems as a difficult problem to formulate this as a scientific inquiry. Taking in account the high diagnostic significance of this feature, the experiments performed aim to determine the smoothness/irregularity of the LMs (borders) in a piecewise manner.

3.1. SMOOTHNESS/IRREGULARITY CHARACTERIZATION

This characterization method includes elements of theory of planar geometry; low pass filtering and one dimensional (1D) signal processing. The latter includes frequency and statistical analysis of the contour of question having in mind that this contour is a non-monotonic 1D signal.

The overall flow of the algorithm for curve segmentation, followed by analysis and classification of the segments is illustrated in Figure 4.

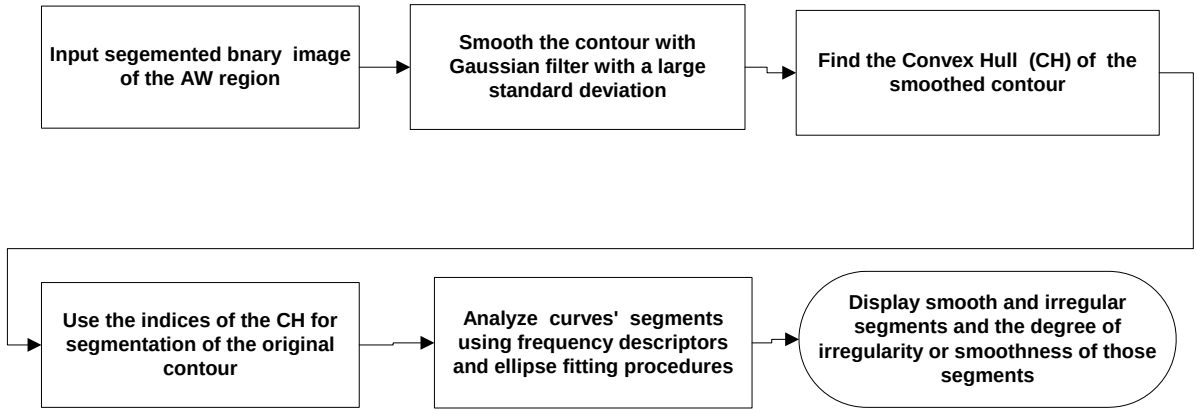

Figure 4. Algorithm for selection of segments and analysis of these segments of a closed contour for smoothness/ irregularity.

The algorithm starts with an automatic election of the boundaries of the object (the detected AW lesion) in the MAP segmented binary image. In a case of multiple lesions, the borders of the segmented AW lesions are extracted automatically and each border is evaluated for its center of a mass $C(\alpha_0, \beta_0)$, circularity $\Psi(x_\alpha, x_\beta)$ and area $\mathrm{A}(x_\alpha, y_\beta)$. This information is used in the overall AW lesion classification.

One of the difficult and challenging facts is that the consecutive points on the curve that belongs to the closed contour are not a monotonic function, which makes the mathematical analysis unconventional and it affects the approach for the meaningful segmentation of the curve. It is difficult to define which points will be the landmarks of the curve for segmentation and smoothness evaluation (Figure 1, 2 and 5). The task's uncertainty can yield unlimited solutions. One way to solve this problem is to find the inflection points of the curve, aka – the curve's 'dominant points.' The 'dominant points' are associated with the second derivative of curvature of the contour [12]. The dominant points are detected by the zerocrossing (the changes in sign) via double differentiation, both on horizontal and vertical axes. This

is not a complex task, but for a curve with many local variations these points can happen to be very near, hence an oversegmentation of the contour in such case occurs.

To increase the distance between the dominant points the contour is smoothed prior the double differentiation. The parameters of the smoothing filter applied before the double differentiation is crucial for the scale-space filtering and density of the dominant points [14]. The smoothing adds an uncertainty because of the low pass filtering, followed by a differentiation that shifts the edges and a significant displacement of the 'true' inflection might occur. Hence, the zerocrossing detection for the segmentation of the curve is dependent on the cutoff frequency of the low pass filter used in scale-space filtering and the size and resolution of the image.

One way to avoid this dependency is to use permanent landmarks of the closed contours. A landmark of a closed curve is the Convex Hull (CH) of curve. The convex hull of a set of points is the intersection of all convex sets which contain the points[13]. If the curve has very frequent curvature changes, the points of the convex hull of the original curve can be too dense. An observation was made that after smoothing the original curve, the convex hull points of the smoothed curve are subset (or at least they correspond) to the CH of the original contour, and the landmark appear sparse. A smoothing with a Gaussian filter with large standard deviation is performed on the original curve and the points of the CH are selected. Example that demonstrates this is shown on Figure 5, where the original contour is displayed in red, the smoothed contour is in blue and the CH points are in yellow and black.

A curve segment which is an object of smoothness analysis is situated between two consecutive points that belong to the CH. Preliminary experiments with different contours with various smoothness confirms the usefulness and the constancy of the segmentation if the CH points approach is used.

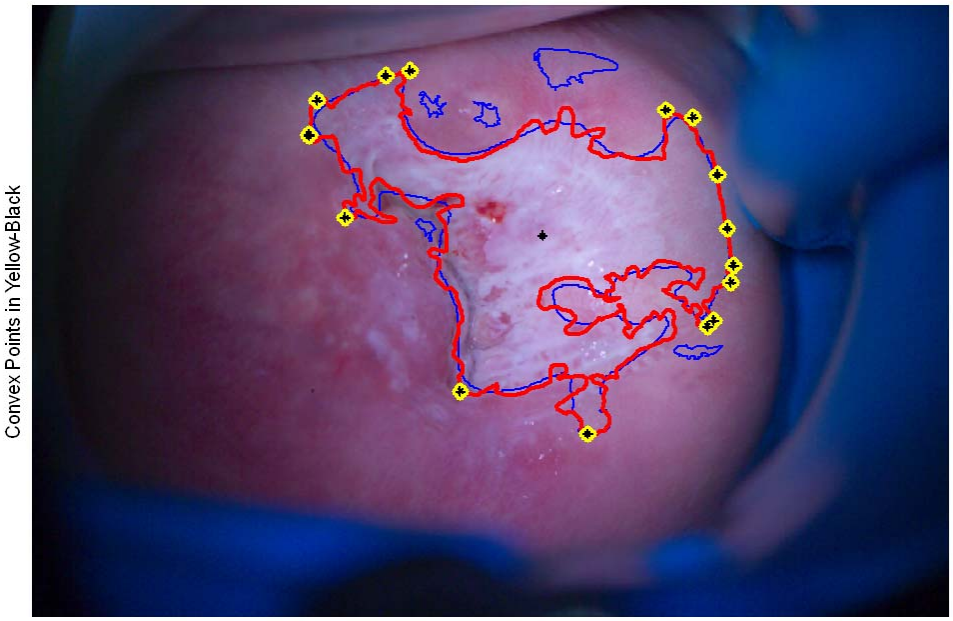

Figure 5. Red delineation shows the original curve enclosing the central part of the AW lesion. The blue curve is the smoothed contour. The black points on yellow background are the points that correspond to the indices of the convex hull of the smoothed contour.

The segmented curves are assessed for smoothness/irregularity in two ways: the first way is by using Frequency Descriptors (FDs) of the smoothed segments vs the segments of the original curve. The second way is by comparing the segments with the corresponding elliptic arc via fitting an ellipse. If both criteria agree on 'smoothness' or 'irregularity' the contour section is classified as 'smooth'.

The FD method is used because this method is independent from the length of the curve segment and it yields an objective comparison between the adjacent segments from the smoothed curve and the irregular segments from the original curve that have the same length and adjacent positions on the AW lesion. The FDs are calculated both for the smoothed curve segments, the original curve segments and the differences in these descriptors are evaluated piecewise (Figure 6).

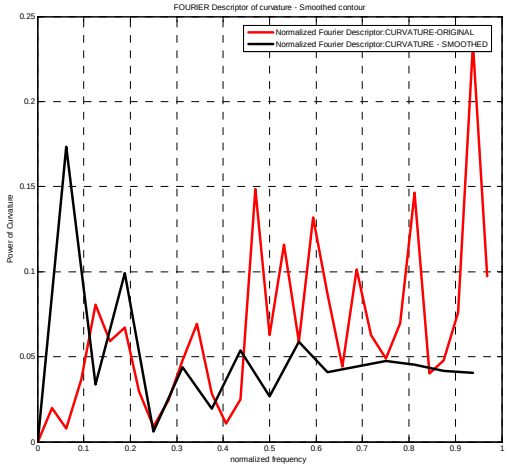

Figure 6. Fourier descriptors of curvature and angularity of margins for smooth (black) and irregular (red) curvature segments.

The FDs are calculated using the Fast Fourier Transform (FFT) for the curvature and angular changes within the segments, substituting the zero frequency value, which contains information of the constancy of the segment. The upper frequencies of the smoothed curve and the original curve segments will differ significantly when the original segment is irregular. This method is not quite robust, so an additional criteria based on the total residual error resulting from an ellipse fitting procedure is added to increase the robustness of the classification.

The ellipse fitting method uses an optimization procedure to fit an ellipse into a set of points using the ellipse-specific fitting algorithm proposed by Halif and Flusser [15] and Hanselman [16]. It's a numerically stable non-iterative algorithm for fitting an ellipse to a set of data points and reformulating the fitting task to optimize this using a quadratic constraint. The Residual Error (RE) process of calculation and arc fitting is depicted in Figure 7. The RE is also normalized by the number of points in the arc (curve segment). As the fitting is equally successful both for smooth and irregular segments, the RE achieved by the 'best fit' for the smooth segments differs from the residual error for the irregular segments and this difference is used as a threshold for the combined classification criteria. The smooth segments are characterized by small residual error (lower threshold), while the irregular has RE two to fivefold more than the RE of the smooth segments (higher threshold).

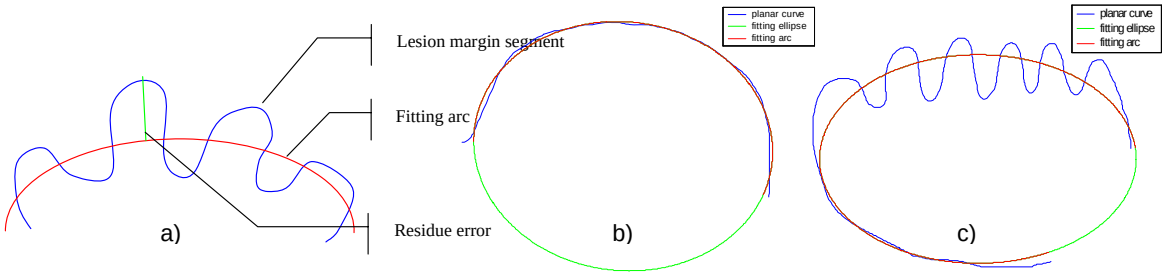

Figure 7. a) Arc fitting (red) of lesion margin border (blue) and the residual error at each point of the curve segment; b) smooth lesion border (blue) fitted well with an ellipse- the residual error is small; c) irregular lesion border (blue) fitted with large residual error.

4. DATA

All colposcopic digital images were acquired during the clinical trials in Lima, Peru in a collaborative project organized by STI-Medical Systems in 2005. Each RGB image has a size of 4500 by 3000 pixels. The images are true color images, originally acquired in a DCR (Kodak – specific format) and later converted in Tagged Image File Format (TIFF). Due to the large memory demand for the preprocessing of such large images; as size-reduction preprocessing procedure was performed. The reduction of the size of the images is achieved by low pass filtering, followed by downsampling by two repeated twice. The lowpass filter has a normalized cutoff frequency of 0.5π and the taps of the filter are normalized in such a way as the DC gain is 0 dB, thus the color of the image remains the same. After this reduction size procedure the images used in the current experiments have the size 1125 by 750 pixels.

The images were taken with a custom-built (research prototype) digital colposcope with two Kodak cameras (left and right). The data was acquired, stored and processed using PC. Two images (left and right) were taken during the colposcopy examination of a patient simultaneously. Between the acquisitions of each image a time interval was allowed (recorded) with intention to study the effect of the temporal changes of the acetic acid. The analysis on both images taken from the left and the right camera is performed in the experiments described here. Because of the positioning of the cameras the two images taken from the same patient differ in illumination conditions and anatomical landmark positions. The preliminary experiments described here were performed on ten images and five patients in total.

5. RESULTS

In this section we present the results to illustrate the segmentation of the AW lesions in Subsection 5.1. The results of the segmentation of the contour, followed by piecewise characterization of segments' smoothness/irregularity are illustrated in the figures in Subsection 5.2. The AW detection followed closely the initial phase for the estimation of the parameters of the regions, followed by the classification the pixels using the 'between-cluster' distance measure introduced in Section 2. The training and testing of the method was using Matlab simulation. For the training phase we used five annotated images where well-defined AW lesions are present. For the testing phase we used ten images, both for detection of the AW lesions and for contour characterization. During the testing we found out that each image contains up to seven lesions that yields the same number of closed boundaries.

5.1. AUTOMATED ACETOWHITE DETECTION VIA MAP ALGORITHM

An automated AW lesion detection using previously described MAP algorithm was performed to obtain a binary image with at least one segmented object. These objects yielded boundaries subjected to further analysis. Original images with

an AW legion located centrally with its automatically delineated red contour are illustrated in Figure 8 a) , c) e) and g). Their adjacent segmentations as binary images are shown in Figure 8 b) , d) f) and h).

Analyzing the results of the AW detection, depicted in Figure 8 we can conclude that the segmentation is consistent with the location of the AW regions for the two patients depicted on Figure 8. Figure 8 a) to d) represents a display of the left and right images acquired from one patient, Figure 8 e) to h) - for the second patient. A large percentage of the segmented AW area corresponds to the actual AW lesions visible of the cervical image. While large variations in illumination can cause a shift in the estimates from the first part of the MAP algorithm it seems that because the images were taken under controlled illumination, these variations are minimal. Examples of AW lesion detection with borders in red are illustrated in Figure 8. The images on the upper row show delineations of a solid 'flat white' AW region with strong internal demarcation borders, pointing toward the cervical canal (CC). The small contour located at 12 o'clock on the cervix contains a diffusive AW in both RGB images, depicted on the first row. The large AW lesion detected centrally contains 'shiny' AW with a strong internal demarcation.

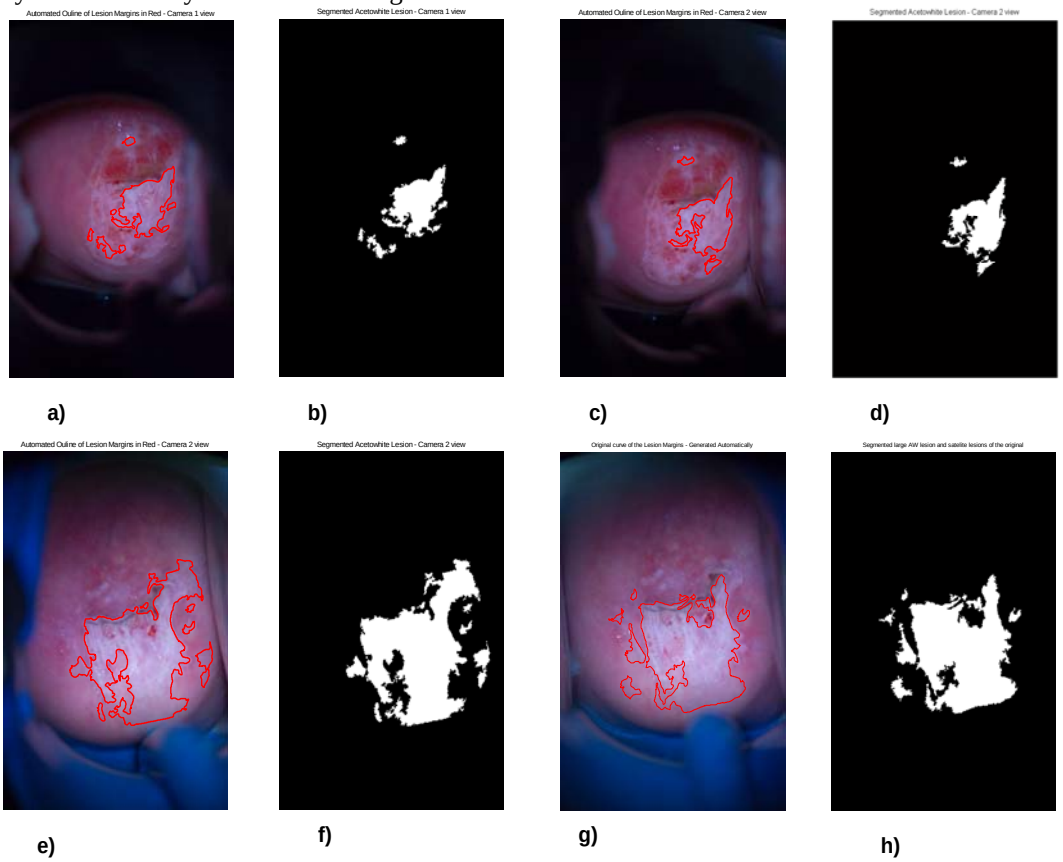

Figure 8. Results of an automated detection achieved using MAP algorithm from two patients (upper and lower row). a) and e) are the RGB colposcopic image taken from the left camera, b) and f) are the adjacent AW binary images; c) and g) are the RGB images taken from the right camera; d) and h) are the adjacent binary image. The red contour outlines the original border of the lesion. There are one central lesion and several multiple lesions of the cervical image.

Similarly to that, the RGB images on the lower row (Figure 8 e) and g)) demonstrates a successful detection of the large AW lesion, located centrally. Small and very diffusive AW areas are detected as neighboring lesions located next to the central AW lesion with 'shiny' acetowhitening that also has a variety of textures.

5.2. SMOOTHNESS/IRREGULARITY OF LESION MARGINS

In this section the results from the LM characterization are displayed in Figure 9 and 10. These results are quite encouraging. The results represented as images and their corresponding contour segmentations and color codes are organized in two rows for each patient. The upper row contains left camera images; the lower row contains the images from the right camera. Images a) and e) represent the original AW lesions detected by MAP algorithm and their boundaries are automatically delineated in red. Images b) and f) represent the smoothed curve on the central lesion only, delineated in blue. Images c) and g) represent the smoothed boundary of the central lesion (blue), the lesions in red and the yellow/black points of the CH, where the curve is segmented for analysis. Images d) and h) are the results of the piecewise smoothness analysis. The yellow lines are the smooth segments that indicate at least CIN II. Colposcopically, the patient was diagnosed with CIN III. Thus the colposcopic diagnosis is the same as the LM characterization yields. Therefore, the results show a success both for the MAP detection algorithm and the LMs segments' characterization of smoothness. Image Figure 9d) shows a very successful characterization of the smooth portions of the LMs, while in a Figure 9 h) shows only the lower lesion margin that was classified successfully as smooth. Perceptually the upper yellow portion of the lesion on Figure 9 h) is partially smooth.

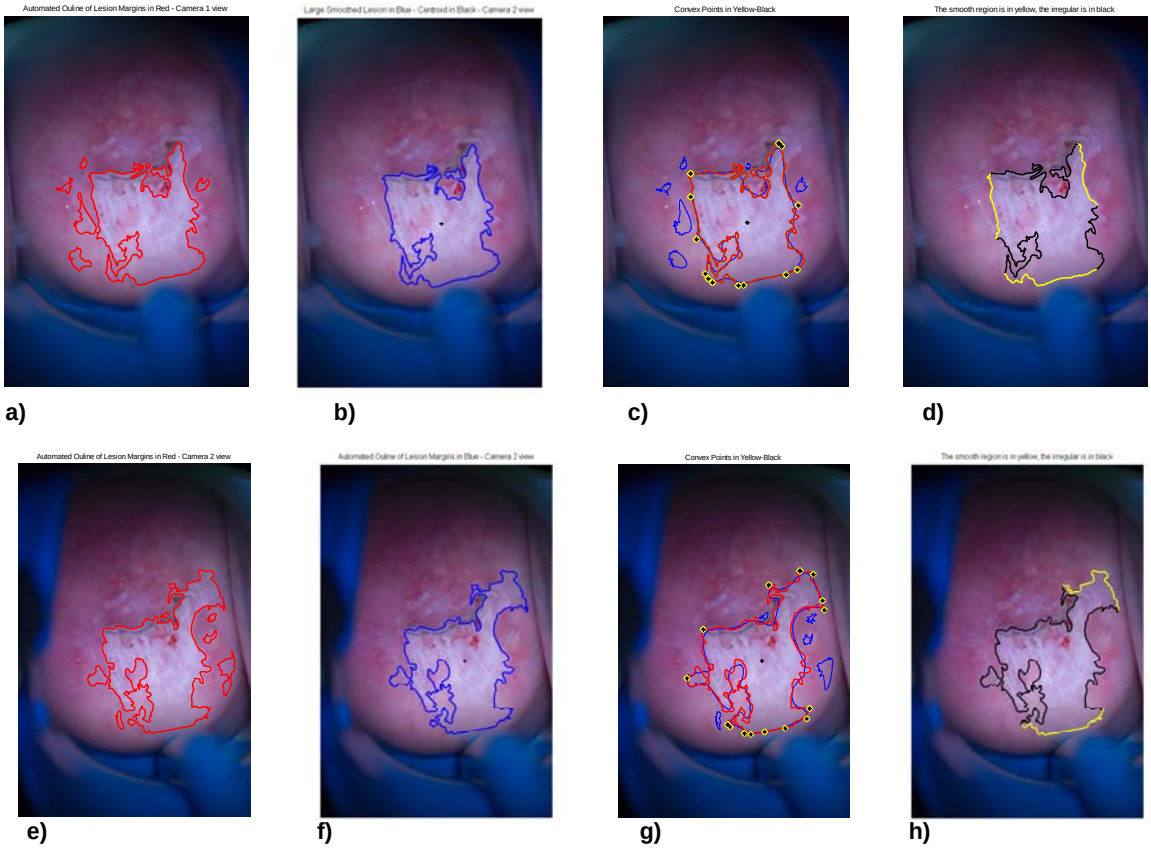

Figure 9. Illustrates the LM characterization of a large AW lesion. Upper row contains the images from the left camera; Lower row contains the images from the right camera. Images a) and e) represent the original AW lesions detected by MAP algorithm and their boundaries automatically delineated in red. Images b) and f) represent the smoothed curve on the central lesion only, delineated in blue. Images c) and g) represent the smoothed boundary of the central lesion (blue), the lesions in red and the yellow/black points for the segmentation of the curve into sections for analysis. Images d) and h) are the results of the piecewise smoothness analysis. The yellow lines are the smooth segments that indicate at least CIN II. Colposcopically the patient was diagnosed with CIN III.

Similarly to this, Figure 10 a) to h) show the success in MAP and LM characterization. Figures 10 d) and h) show right classification of the smooth section with exception of the classification on the lower rim of the LM in image h), which appears to be irregular, rather than smooth. In all of the cases the presence of smooth pieces indicate CIN III grade. Thus, the automated diagnosis is the same as the colposcopic diagnosis.

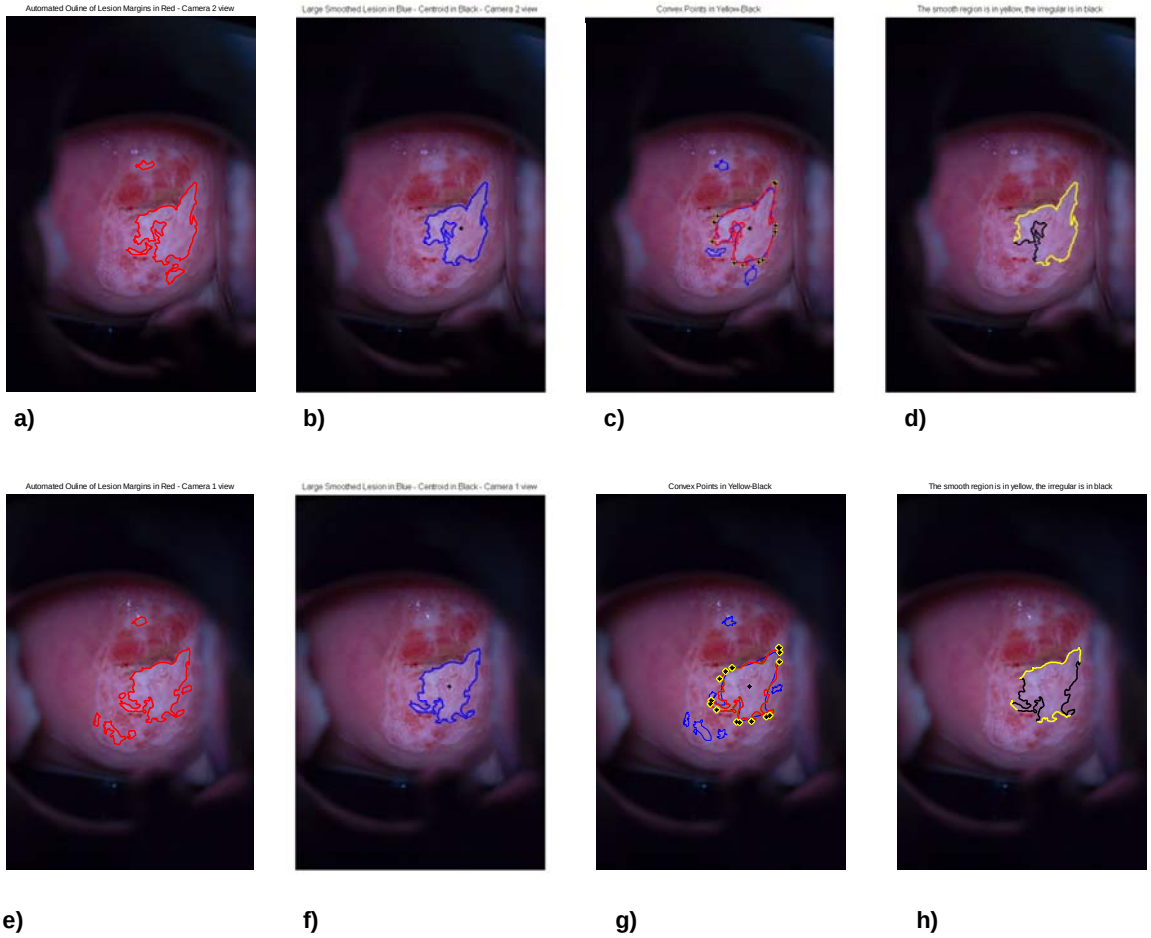

Figure 10. Illustrates the LM characterization of a large AW lesion. Upper row contains the images from the left camera; Lower row contains the images from the right camera. Images a) and e) represent the original AW lesions detected by MAP algorithm and their boundaries automatically delineated in red. Images b) and f) represent the smoothed curve on the central lesion only, delineated in blue. Images c) and g) represent the smoothed boundary of the central lesion (blue), the lesions in red and the yellow/black points for the segmentation of the curve into sections for analysis. Images d) and h) are the results of the piecewise smoothness analysis. The yellow lines are the smooth segments that indicate grade higher than CIN II. Colposcopically the patient was diagnosed with CIN III.

6. DISCUSSIONS AND CONCLUSIONS

In this paper we discussed the importance of detecting the AW lesions and evaluating their borders in terms of their smoothness or irregularity of the LMs segments. The LMs characterization can be important element of a computerized support for cervical cancer diagnosis. In addition, the importance and the methods for detecting and characterizing the AW lesions were also discussed. We suggested a novel method for color-based detection of the AW lesions. We suggested a novel method for meaningful curve segmentation and smoothness/irregularity characterization of the segments of this curve. In addition to this we reported encouraging result regarding AW detection and border analysis.

Demonstrated were novel methods such as MAP color based segmentation and curve analysis with FDs and ellipse fitting techniques. As the clinical data was acquired recently, the test results are preliminary both for the MAP algorithm and the LMs algorithm.

The results of the studies of the LMs are promising as we succeeded to correctly classify the disease in comparison with the colposcopic diagnosis. Regardless of the limited number of the tested images the developed methods are robust and stable. The algorithm for irregular/smooth segments for LMs characterization needs refinements to establish better classification criteria based on the FDs and ellipse fitting methods. This also is necessary to be tested extensively on the full data set. Furthermore, the LMs characterization can be extended to analyze for diffusivity/demarcation or lesions' location. Each one of these studies requires its own experimental testing and verification. We continue further the collaborative work with gynecologists and colposcopists for algorithm-wise development in order to achieve better diagnosis.

7. ACKNOWLEDGMENTS

The authors of this paper would like to thank the Instituto Especializado de Enfermedades Neoplasicas (INEN), Lima, Peru and especially Dr. Carlos Santos and Ms. Carmen Céspedes Cevasco at the Department of Obstetrics and Gynecology in Lima, Peru for the great collaborative work with STI – Medical Systems. Our gratitude is extended to Dr. Marcelo Soto-Thompson and Dr. Jody Oyama for organizing the clinical trials, patient relations and data acquisition. The authors thank Dr. Daron Ferris for his valuable advice and sharing his extensive colposcopic expertise in the clinical aspects of the described algorithms.

8. REFERENCES

1. E. L. Franco, E. Duarte-Franco, A. Ferenczy, Cervical Cancer: Epidemiology, Prevention, and the Role of Human Papillomavirus Infection, *CMAJ*, Vol. **16**, pp.1017-1025, 2001.
2. V. Beral and M. Booth, Precision of Cervical Cancer Incidence and Mortality in England and Wales, *Lancet*, Vol. **1**, pp. 495, 1986.
3. M. L. Steward and Y. C. Collins, Strength of Correlations between Colposcopic Impressions and Biopsy Histology, *Gynecol. Oncol.*, Vol. **89**, pp. 424-428, 2003.
4. V. Van Raad and A. Bradley, Emerging Technologies, Signal Processing and Statistical Methods for Screening of Cervical Cancer in vivo – Are they good candidates for Cervical Screening?, *IEE Proceedings of Advances in Biomed. Signal and Information Processing*, Vol. **2**, pp. 210-217, 2004.
5. W. Li, V. Van Raad, et al., Computer-aided Diagnosis (CAD) for Cervical Cancer Screening and Diagnosis: a New System Design in Medical Image Processing, *Lecture Notes in Computer Science, CVBIA 2005*, pp. 240-250, 2005.
6. C. Balas, G. Themelis, E. Prokopakis, I. Orfanudaki, E. Koumanatakis, E. Helidonis, In vivo Detection and Staging of Epithelial Dysplasias and Malignancies Based on the Quantitative Assessment of Acetic Acid-tissue Interactions Kinetics, *Journal of Photochemistry and Photobiology*, ***B:** Biology*, Vol. **53**, pp.153-159, 1999.
7. R. Reid, and P. Scalzi, Genital Warts and Cervical Cancer. VII. An Improved Colposcopic Index for Differentiating Benign Papillomaviral Infections from High-Grade Cervical Intraepithelial Neoplasia, *Am. J. Obstet. Gynecol.*, Vol. **153**(6), pp. 611-618, 1985.
8. Claude I., Winzenrieth R., Pouletaut P., Boulange J-C, "Contour Features for Colposcopic Image Classification by Artificial Neural Networks," (ICPR'02), Vol. 1, pp.771-774, 16th International Conference on Pattern Recognition, 2002.
9. Claude, I., Pouletaut, P., Huault, S.and Boulanger, J.C., Integrated color and texture tools for colposcopic image segmentation, *ICIP***,** Vol. **2** , pp. 311 -314, 2001.
10. P. J. Disaia, W. T.Creasman, Clinical Gynecologic Oncology, 3rd Ed. The C. V. Mosby Company, Toronto, pp. 1-13, 1989.
11. V. Van Raad, V., Image Analysis and Segmentation of Anatomical Features of Cervix Uteri in Color Space, *Proceedings of 13th European Signal Processing Conference – EUSIPCO*, 2005.
12. T. Lee, D. McLean, M. S. Atkins, Irregularity index: A New Border Irregularity Measure for Cutaneous Melanocytic Lesions, *Med. Image Analysis*, Vol. 7(1), pp. 47-64, 2003.
13. http://www.york.cuny.edu/~wu/jm/Glossary.htm
14. A. P. Witkin, Scale space filtering, *Proceedings of 8th International Joint Conference on Artificial Intelligence*, pp. 1019-1022, Karlsruhe, West Germany, August, 1983.
15. R. Halif and J. Flusser, Numerically Stable Direct Least Squares Fitting of Ellipses, Department of Software Engineering, Charles University, Czech Republic, 2000.
16. D. Hanselman, Stable Direct Least Squares Ellipse Fit, *Matlab Central*.