Time-Resolved Imaging of Cervical Acetowhitening

Philippe Schmid-Saugeon, Jonathan D. Pitts, Howard B. Kaufman, Alex Zelenchuk, and Diane M. Harper

Abstract

Acetowhitening is a phenomenon that can be observed when applying acetic acid to cervices. The way tissues whiten is observed and rated in *colposcopy* to help diagnose any precancerous changes in the cervix. This paper reports a study designed to measure multiple parameters of the acetowhitening process from a sequence of images captured with an improved *colposcope* after the application of acetic acid. Clinical examination was recorded on video tape in order to allow for the registration of the image data with the biopsy locations for which a histopatology confirmed diagnosis was available. Image processing techniques were used to align, filter, and segment the images of size 640×480 captured every 5 seconds during 5 minutes. A *whiteness signal* was then computed for each biopsy location. After fitting a polynomial-based model to the data and computing different features, a *linear discriminant analysis* was performed to separate high-grade lesions from other types of lesions. In this study we used 26 *high grade* (CIN II-III) pathology-confirmed lesions from 22 unique patients, 16 *low grade* (CIN I) pathology-confirmed lesions from 13 unique patients, 11 lesions evaluated by pathology as *no evidence of disease* (NED) from 10 unique patients, and 5 lesions evaluated by pathology as *metaplasia* (MET) from 5 unique patients. Using two single features, we could achieve a sensitivity and a specificity of 88 % for the separation of CIN II-III from the combination of CIN I, NED, and MET, while for regular colposcopy one can expect a specificity of 69 % for a sensitivity of 85 % [1].

Index Terms

Cervical cancer, colposcopy, acetowhitening, computer-aided diagnosis, image processing

I. Introduction

Approximately 3.5 million women in the United States are diagnosed annually with a cytological abnormality requiring additional follow-up or evaluation. Nearly three million are referred

for a colposcopic examination to help diagnose any precancerous changes that should be treated to prevent invasive cervical cancer. Colposcopy is the final diagnostic examination for a woman with abnormal cytology prior to excising the questionable tissue itself.

Colposcopy of the cervix is an examination with high magnification and bright intense white light under the conditions of a 5 % acetic acid tissue soak. The acetic acid reacts with the cervical tissue producing observable degrees of tissue whitening. This change in tissue appearance is called *acetowhitening* and is used as an indicator for Cervical Intraepithelial Neoplasia (CIN). Cells infected with the human papillomavirus (HPV) also become acetowhite upon the application of 3-5 % acetic acid. No other viral or bacterial infection causes an acetowhitening change.

The mechanism for acetowhitening has not been identified, though a variety of hypotheses have been proposed. Burke suggests, ”acetic acid causes water to leave the cell, after which the cell membrane collapses around the abnormal and large nucleus” [2]. Singer states the ”acetic acid causes the tissues, especially the columnar and abnormal epithelia, to become edematous. The acetic acid seems to be causing the coagulation of epithelial and stromal cytokeratins in a reversible event” [3]. Others put forward the predominant effect in acetowhitening is ”an alteration of the refractive-index structure of the [cell's] nucleus” [4] and that in addition to changes in optical back-scattering properties of cells, acetic acid may also increase the depolarization of reflected light from intra-nuclear structures [5].

There is universal agreement, however, that the use of acetic acid and observation of any resulting acetowhitening, is necessary for a colposcopic examination to be deemed satisfactory. Colposcopists tend to grade the intensity of acetic acid whitening subjectively in terms of white, whiter, and whitest. “The degree of whitening is subjectively used as an aid to identify appropriate tissue to biopsy for histological examination. In addition to the intensity of the whiteness, the speed at which it occurs, its duration, and the rapidity of its disappearance also correlate with the underlying histopathology” [2]. By measuring the maximum intensity of acetic acid whitening and the rates of increase and decay of this intensity using direct reflectance throughout the visible and near infrared spectral regions, more quantitative assessments can be made which may improve the early detection of cervical cancer and reliable monitoring of pre-cancerous conditions.

CIN is graded according to severity, with CIN I being the mildest form, which often regresses back to normal tissue without treatment, to CIN II-III, the more severe of the pre-cancerous

lesions. Clinical practice regarding CIN II-III is consistent. When found, CIN II-III is treated with tissue excision, usually with a procedure referred to as Loop Electrical Excise Procedure (LEEP). Unfortunately, colposcopy has tremendous intra-observer variability when it comes to identifying CIN II-III lesions. Hopman et. al., published the results of one study which found the level of agreement among 23 experienced colposcopists to identify CIN III lesions to be 76.9 % [6]. The level of experience of the colposcopist also influences the accuracy of the exam. For a graduating gynecology resident, the accuracy (sensitivity) of colposcopy is between 20-35 % [2]. For expert, experienced colposcopists the best accuracy is 75 % [2]. In general colposcopy is imprecise, although useful in estimating lesion grade [7].

Balas et al [8], [9] have tried to characterize the acetowhitening of the cervix to indicate high-grade pathology, but did not use biopsy-registered histology to develop their algorithm. Very few patients were used in their study, and for illustration purpose only. Their claims were not proven but only suggested by a few observations. Finally, they gave almost no details about image and video processing techniques. It is unclear how they handled image alignment, for example.

The purpose of this study is to work with objective measures to characterize the CIN II-III precancerous condition of the cervix. Improvements upon the subjective nature of colposcopy can help identify those women truly at risk for cervical cancer and prevent overtreatment of those women who are not at risk for cervical cancer. In this paper the AW process is described mathematically. This study cannot be used to suggest a better clinical practice, for which more data and better statistics, i.e. better histopathology, better correlation with histopathology, and a better classification system, would be necessary.

This paper is organized as follows: Section II describes the colposcopic workstation and the way image sequences are acquired. Image alignment is treated in Section III, temporal and spatial filtering in Section IV, and the segmentation of sequences of acetowhitening images in Section V. Data fitting and the extraction of whitening features are explained in Section VI, and the technique used to classify the pathology confirmed lesions is given in Section VII. Finally, results are given in Section VIII and conclusions are drawn in Section IX.

II. MATERIAL

A. Acquisition system

Clinical data was collected using an improved colposcopic workstation. The hardware consists of a standard Carl Zeiss colposcope system (model OPMI-1FC, ZMS-506-II), illuminated by a fiber optic-coupled 15 volt/150 watt halogen lamp with 16x eye binoculars. The net magnification of the microscope is 200 $\times$. A dual port video splitter is used to mount a Sony 3-CCD RGB camera (model DXC-390). Images from the color CCD were actively displayed on the workstation's flat panel monitor. An optical filter assembly consisting of a fixed linear polarizer and a second rotatable linear polarizer was placed in the instrument's optical path. When the filters were cross polarized unwanted glare from the surface of the cervix was reduced. This cross-polarized state was used for all data collection. A video tape recorder (JVC model KZV10) was also attached to an auxiliary output of the Sony 3-CCD color camera to aid in documenting biopsy locations taken during the colposcopic examinations. A video capture board (Imagenation model PXC-200) was used in the workstation's computer.

B. Patients & Data

Women undergoing routine diagnostic colposcopy at the Dartmouth-Hitchcock Medical Center (DHMC), Hanover, NH, were asked to participate in this study and indicated their willingness to participate by signing an informed consent. Pregnant women, women younger than 18 years of age or older than 65 years of age were excluded from the study. Informed consent was obtained from the patient by the study's DHMC clinical coordinator. Per the IRB-approved clinical protocol, all women underwent a routine colposcopic procedure and biopsies were acquired only as needed per standard clinical practice. The cervix was visualized with the study's colposcopic workstation and 0.7 ml of 5 % acetic acid solution was applied uniformly across the surface with a mucosal atomizer device (Wolfe Troy Medical, Model MAD200). Digital images were recorded at the 2.6 $\times$ magnification setting giving a panoramic view of the entire cervical field

Images of size 640×480 were taken at $N = 60$ time steps after the application of acetic acid. The time interval between successive images was 5 seconds, giving a total recording time of 5 minutes. Images were aligned as explained later in Section III. The procedure videotape was reviewed to identify the biopsy locations on the surface of the cervix. For accurate determination,

the videotaped procedure was reviewed next to a computer monitor displaying a still image taken from the 5-minute time-sequence. This *reference image* was chosen at a peak whitening stage. It was annotated in a painting software to mark the areas of the cervix congruent with the biopsy location observed in the video. *Whiteness data* $\mathbf{w}(i,j;n)$ was extracted at each pixel location $(i,j) \in \mathbb{I} = \{0,\ldots,r-1\} \times \{0,\ldots,c-1\}$ and each time step $n \in \mathbb{T} = \{n_0,\ldots,n_{N-1}\}$, where $r = 480$ is the number of rows and $c = 640$ is the number of columns. $\mathbb{I}$ and $\mathbb{T}$ are the image and time domains, respectively. This whiteness data was then averaged for regions matching biopsy locations. These regions were obtained by means of the image sequence segmentation techniques detailed in Section V. For a region with label k, the average whiteness signal is given by:

$$\mu_{\mathbf{w}}(k;t) = \frac{1}{N_k} \sum_{(i,j)\in\mathbb{I}_k} \mathbf{w}(i,j;t) \tag{1}$$

where $\mathbb{I}_k \subset \mathbb{I}$ is the set of all pixels that belong to the region k and N_k is the size of $\mathbb{I}_k$.

To construct whiteness signals, a whiteness metric must be defined, i.e. an adequate RGB to *intensity* transformation must be chosen. With *adequate* we mean the metric that allows for the best discrimination between different tissue classes when it comes to classify the collected data. We went through the different processing steps, from the image alignment to the final classification, using different metrics, such as each single R, G, or B channel, intensity CCIR [10], luminance L^* [11], saturation in the $L^*u^*v^*$ [11] color space, saturation in the Munsel system [12], etc. We did not find that any metric outperformed the others. A slightly better result was obtained with the intensity CCIR, which is the one that was used to obtain the results presented in this paper. It is given by [10]:

$$I = 0.299R + 0.587G + 0.114B \tag{2}$$

Since we are interested in the way the cervix whitens after the application of acetic acid and not in the absolute intensity of the cervix surface, all whiteness signals were background subtracted. With background we mean the intensity of the cervix before the acid is applied. Therefore, the following transformation was applied to all signals:

$$\mathbf{w}(i,j;n) \rightarrow \mathbf{w}(i,j;n) - \mathbf{w}(i,j;n_0), \forall n \in \mathbb{T} \tag{3}$$

The different image and video processing techniques leading to the biopsy-registered whitening signals are detailed in the next sections.

III. IMAGE ALIGNMENT

Before we can extract the whiteness signals $\mathbf{w}(i,j;n)$, we need to align the images in each sequence. The transformation of the coordinate system between successive image pairs was assumed to be translation only. In order to reduce as much as possible the distortion due to local deformations, images were aligned in two phases:

1) global alignment,
2) local alignment.

The former corrects large shifts due to operator manipulations or patient movements. The latter corrects local motion due to muscle contractions mainly. Local alignment was performed where biopsies were taken.

To compute the motion vector between two sub-images, we used *normalized cross-correlation* [12]. In order to rely only on the structural information contained in the images that had to be aligned, the alignment was based on the *Laplacian-of-Gaussian* (LoG) [12] images. The normalized cross-correlation was computed between an $N_1 \times N_1$ centered window taken from the first LoG image, and an $N_2 \times N_2$ centered window taken from the second LoG image, with $N_1 > N_2$. In our implementation, $N_1 = 256$ and $N_2 = 128$. Only the top-left area of size $N_1 - N_2 + 1$ of the result was searched for a maximum, whose location defines the motion vector. The normalized cross-correlation is given by [12]:

$$\mathcal{CC}_{\text{norm}} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{I}_1)\mathcal{F}^{*}(\mathbf{I}_2))/(\mathbf{E}_1 E_2) \tag{4}$$

$$\mathbf{E}_1 = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{I}_1^2)\mathcal{F}^{*}(\mathbf{I}_{\text{ones}})) \tag{5}$$

$$E_2 = \sum \mathbf{I}_2^2 \,, \tag{6}$$

where $\mathcal{F}(\cdot)$ and $\mathcal{F}^{-1}(\cdot)$ are the Fourier transform and its inverse, respectively, $\mathbf{E}_1$ and E_2 are the neighborhood energy of $\mathbf{I}_1$ and the total energy of $\mathbf{I}_2$, respectively, and $\mathbf{I}_{\text{ones}}$ is a matrix of size $N_1 \times N_1$ having ones in the top-left area of size $N_2 \times N_2$ and zeros elsewhere. Note that $\mathbf{E}_1$ is a matrix of size $N_1 \times N_1$, while E_2 is a scalar. In order to compute the cross-correlation, the region of support of $\mathbf{I}_2$ is extended to $N_1 \times N_1$ by padding with zeros.

Since full automation was not a requirement in this phase, a *graphical user interface* (GUI) was written in Matlab in order to validate the local alignment and if necessary to correct it. This was rarely the case.

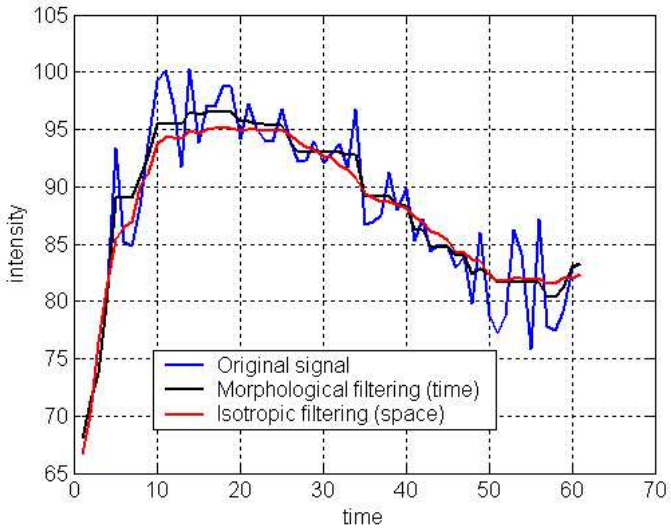

Fig. 1. Temporal and spatial filtering of acetowhitening signal.

IV. Temporal and spatial filtering

The influence of noise due to acquisition hardware, to changing lighting conditions, and misaligned image pairs must be reduced as much as possible. A first filter was applied to the time axis, individually for each pixel coordinate. We used the following morphological filter:

$$\mathbf{w}(t) \odot b = \frac{1}{2}\left[(\mathbf{w} \circ b) \bullet b + (\mathbf{w} \bullet b) \circ b\right] \tag{7}$$

where b is called the structuring element, $\circ$ is the opening operator, and $\bullet$ is the closing operator [13]. The structuring element had the shape of a half circle.

A Gaussian low-pass filter was applied to the spatial domain. It was implemented in an iterative fashion using a diffusion equation [14]. An example of filtered signal is shown in Fig. 1.

V. Segmentation of acetowhitening sequences

A. Introduction

The objective of image sequence segmentation was to outline lesion candidates for which biopsies were taken. In this section we briefly review the different techniques used in this study.

The common background of all techniques is the *similarity measure* between two whiteness signals and the segmentation into a *grain image*, as explained in § V-B and V-C, respectively.

B. Similarity measure

In order to measure the similarity, or inversely the difference between two whiteness curves, we had to define a function called hereafter *fitting function*. The main problem that arises when processing the raw whiteness data is its high dimension. Sixty-dimensional vectors are typical whiteness signals, and it is out of question to use all data to segment the image plane. It is necessary to reduce the data dimension, down to a one-dimensional vector if possible. Such a scalar can be obtained either by quantifying a leading characteristic of the curve or, for example, by taking the inner product of both vectors. We define the following fitting function:

$$\varphi\left(\mathbf{x}_1, \mathbf{x}_2\right) = \frac{\left\langle \mathbf{w}\left(\mathbf{x}_1; t\right), \mathbf{w}\left(\mathbf{x}_2; t\right)\right\rangle}{\max\left(\Omega\left(\mathbf{x}_1\right), \Omega\left(\mathbf{x}_2\right)\right)} \tag{8}$$

where $\mathbf{x}_1$ and $\mathbf{x}$ are two pixel coordinates, and $\Omega\left(\mathbf{x}_1\right) = \left\langle \mathbf{w}\left(\mathbf{x}_1; t\right), \mathbf{w}\left(\mathbf{x}_1; t\right)\right\rangle$ is the energy of the whiteness signal at location $\mathbf{x}_1$. Note that we have

$$0 \leq \varphi\left(\mathbf{x}_1, \mathbf{x}_2\right) \leq 1 \tag{9}$$

since all whiteness signals are supposed to be positive-valued signals. Note that the inner-products are normalized by the highest signal energy and not by the square root of the product of both energies. The reason is that scaled versions of a whiteness signal must not be considered as being identical.

C. Image grains

In order to reduce the amount of data that has to be processed and to improve the signal-to-noise ratio of the whiteness signals, the image plane can be pre-segmented into *grains*. Those small regions are the result of a coarse segmentation. Depending on the sequence, the mean grain surface is around 30 pixels, but they can be as small as a few pixels or as large as a few hundred pixels.

To achieve the initial segmentation, we chose a very quick and efficient technique for this type of tasks, namely the watershed transform [15]. To make it short, an image is considered as a three-dimensional surface with a hole at each local minimum. Flooding is then simulated by

progressively increasing the *water* level, starting at a level equal to the global minimum. Each time the water level reaches a new hole, it starts flooding the neighborhood, called catchment basin. The flooded area gets a new label. When the water level is such that two neighboring areas meet, then the meeting points form what is called the watershed. A fast implementation of the watershed transform can be found in [16].

The watershed transform is applied to a gradient image, which in our case is computed as explained in the next paragraph.

D. Gradient image

The gradient image of an intensity image is the approximation of the amplitude of the local gradient at every pixel location. In the case of whiteness signals, we can use the similarity function defined by Eq. 8 to obtain a correlate for the difference in whitening. We compute the fitting value $\varphi(i,j;i_0,j_0)$ between a pixel (i_0,j_0) and all its neighbors $(i,j) \in \mathbb{N}_{(i_0,j_0)}$, where

$$\mathbb{N}_{(i_0,j_0)} = \{(i_0-1,j_0),(i_0,j_0-1),(i_0+1,j_0),(i_0,j_0+1)\} \tag{10}$$

Since the best fit should correspond to a null gradient, i.e. the neighborhood is homogeneous, we compute the following function:

$$\varphi_\nabla(i,j;i_0,j_0) = \delta\frac{1-\varphi(i,j;i_0,j_0)}{1+\varphi(i,j;i_0,j_0)} \tag{11}$$

where $\varphi_\nabla(i,j;i_0,j_0) \in (-\infty,\infty)$. The sign of $\varphi_\nabla(i,j;i_0,j_0)$ is given by:

$$\delta = \begin{cases} 1 & \text{if } \Omega(i_0,j_0) \geq \Omega(i,j) \\ -1 & \text{if } \Omega(i_0,j_0) < \Omega(i,j) \end{cases} \tag{12}$$

Finally, the derivatives are approximated as the mean of the forward and backward differences:

$$\frac{d}{dx}\mathbf{w}(i_0,j_0) = \frac{\varphi_\nabla(i_0,j_0-1;i_0,j_0) - \varphi_\nabla(i_0,j_0+1;i_0,j_0)}{2} \tag{13}$$

$$\frac{d}{dy}\mathbf{w}(i_0,j_0) = \frac{\varphi_\nabla(i_0-1,j_0;i_0,j_0) - \varphi_\nabla(i_0+1,j_0;i_0,j_0)}{2} \tag{14}$$

from which we can approximate the norm of the gradient vector. Figures 2 and 3 show two examples of gradient images: the gradient image using the approximations given by Eqs. 13 and 14, and the gradient image obtained for the reference image only, respectively. It is visible from those two images that the former contains more details visible only in some of the images. The glare makes it visible (inside the red circle): it is moving with time.

Fig. 2. Gradient image of sequence.

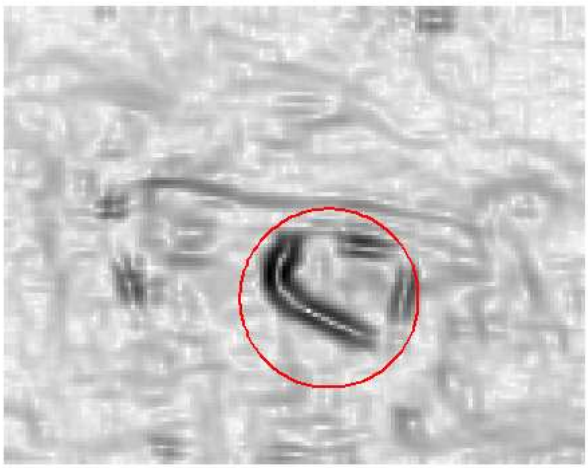

Fig. 3. Gradient image of reference image.

E. Robust region-merging

Region merging is a well known approach to image segmentation [12]. The principle is simple: initially each pixel is a region, and neighboring regions are merged according to a predefined criterion in an iterative fashion. In our case, the criterion is a threshold set on the value provided by the fitting function (see Eq. 8). The segmentation converges to the final result when no pair of neighboring regions meet the merging criterion. In the case of acetowhitening data, we want to merge regions whose whiteness data is similar. The fitting function will therefore measure the similarity between two signals and decide if the similarity is sufficient for merging. Whiteness signals are updated after each merge with Eq. 1.

The objective of the robust region-merging algorithm is to take into account the *homogeneity* of the data inside the different regions. While the region-merging algorithm relies only on the mean whiteness signal of each region to decide subsequent merging, this algorithm allows for controlling the maximum variability allowed inside each region. More specifically, the variance signal associated to each region is computed:

$$\begin{aligned}\sigma^2_{\mathbf{w}}(k;n) &= \frac{1}{N_k}\sum_{(i,j)\in\mathbb{I}_k}\left(\mathbf{w}(i,j;n)-\bar{\mathbf{w}}(k;n)\right)^2 \\ &= \frac{1}{N_k}\sum_{(i,j)\in\mathbb{I}_k}\mathbf{w}^2(i,j;n)-\left(\frac{1}{N_k}\sum_{(i,j)\in\mathbb{I}_k}\mathbf{w}(i,j;n)\right)^2\end{aligned} \tag{15}$$

and the energy of the standard deviation signal is used as merging criterion:

$$\langle\sigma_{\mathbf{w}},\sigma_{\mathbf{w}}\rangle = \sum_{t\in\mathbb{T}}\sigma^2_{\mathbf{w}}(k;n) \tag{16}$$

While the above criterion is used to decide which regions can be merged, the order in which regions are merged is dictated by the increase of variance each merged pair would generate:

$$\Delta_\sigma(k,l) = \sum_{t\in\mathbb{T}}\left(\sigma^2_{\mathbf{w}}(k\cup l;t)-\frac{1}{2}\left[\sigma^2_{\mathbf{w}}(k;t)+\sigma^2_{\mathbf{w}}(l;t)\right]\right) \tag{17}$$

With this approach, if a region can be merged with more than one of its neighbors, it is only merged with the one that increases less the variance. The other neighbors will have to wait until the next iteration to be merged, given that they still meet the fitting criterion with their updated neighbor.

F. Clustering

Another class of segmentation techniques is based on the clustering of data. While numerous different clustering techniques exist, we will focus in this section on the fuzzy c-means technique [17]. We had successfully applied this technique in the past [18] and we relied on this experience to make our choice. The fuzzy c-means clustering technique is an iterative technique where partitions and centroids are successively updated. Data points are related to the partitions by means of a *fuzzy membership value*. The different equations and their derivation can be found in [17].

The distance functional we used in this study was based on the fitting function given by Eq. 8. However, since similar whiteness signals are supposed to be close to each other, we define the following distance function:

$$\|\mathbf{x}_k - \mathbf{v}_i\| = \frac{1}{2}\left(1 - \varphi_{ki}\right) \tag{18}$$

where $\mathbf{x}_k$ is a data point, $\mathbf{v}_i$ is the centroid of the i^{th} cluster, and φ_{ki} is given by Eq. 8. The number of clusters was chosen empirically and was most of the time set to 3.

G. Hierarchical watershed

Another way to merge neighboring grains is to continue working with morphological techniques. To this end, we can use a technique called hierarchical watershed [19]. The idea is quite simple:

1) Segment an image with the watershed transform,
2) Reconstruct the gradient image from the watershed,
3) Return to step 1 until a merging criterion is met.

The core concept here is geodesic reconstruction [20]. This algorithm converges to a unique region if no stopping criterion is used. Such a criterion can simply be a predifined number of iterations, as it was the case in our study.

H. Examples of segmentation results

The different techniques proposed in this report were integrated into a research tool called AWSEGMENT. This graphical user interface written in Matlab gives the user the possibility to segment acetowhitening data and to display intermediate and final results. The objective is (1)

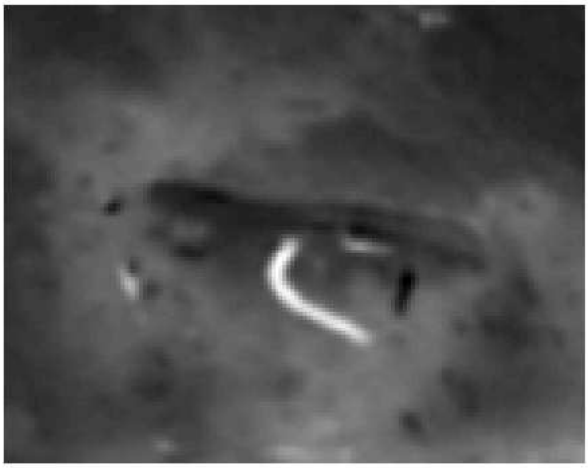

Fig. 4. Background subtracted reference image.

to define which techniques are best adapted to our data and (2) to process the data and to extract the mean curves for each region for which a biopsy was taken.

We have experienced that the best strategy to adopt when processing acetowhitening sequences is a *coarse-to-fine* approach. A first segmentation outlines large whitening regions, called hereafter *background lesions*, which are then considered as regions-of-interest and masked out for additional segmentation. A second segmentation step will outline smaller lesions, called hereafter *foreground lesions*. The final regions will then be combined with the manually drawn biopsy locations.

Figures 4 and 5 show the reference image and the grain image for the same patient, respectively, both filtered and background subtracted. To obtain the grain image, all pixel values inside a region are averaged individually for each time step. The signal-to-noise ratio is therefore increased, while there is sufficient resolution left to recognize relevant regions and possible lesions.

Figures 6 and 7 show one and two iterations of hierarchical watershed, respectively. It can be seen that this approach is very useful to segment larger areas, like large lesions or the full cervix,

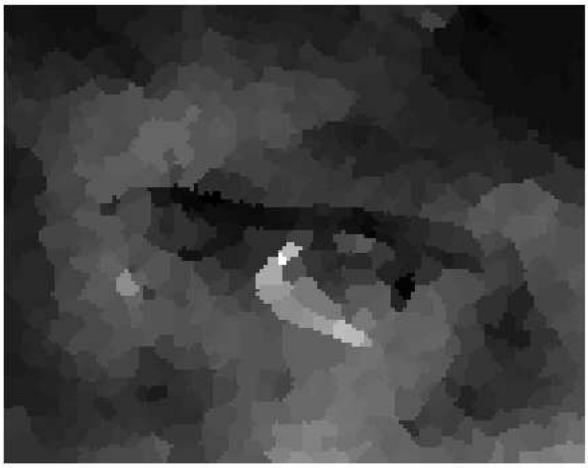

Fig. 5. Grain image obtained after watershed segmentation.

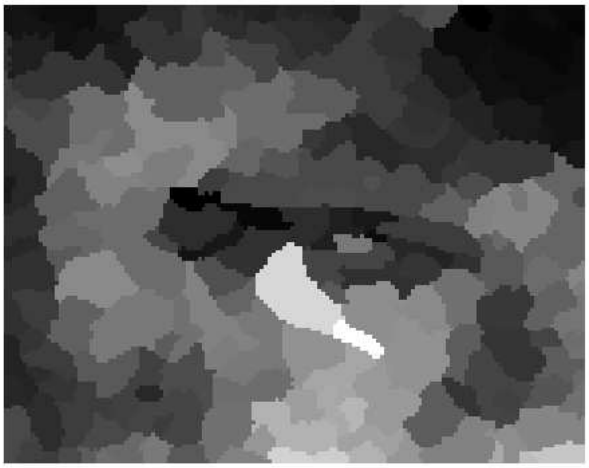

Fig. 6. Segmentation result after one iteration of hierarchical watershed. The original watershed is displayed in Fig. 5.

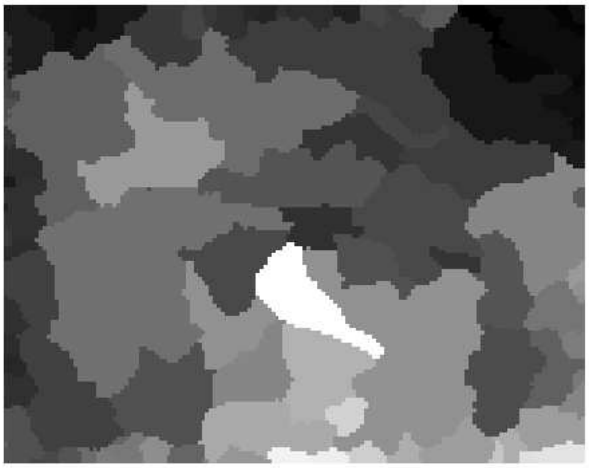

Fig. 7. Segmentation result after one iteration of hierarchical watershed.

and that only a few iterations are necessary to reach the point where regions become too large to correspond to any significant lesion in the image. For our application, we found it sufficient to make a single iteration of the hierarchical watershed (Fig. 6) and to continue merging regions with the robust region-merging technique.

Figure 8 shows the segmentation result obtained with the clustering technique (3 regions). Regions with an area smaller than 16 pixels were removed. Each region has a different label, and pseudo-colors were used only for display purpose in order to improve the contrast between neighboring regions. A boomerang shaped background lesion can be seen in Fig 8. It is segmented again with the clustering technique. The final mask is displayed in Fig. 9 and the mean curves computed for each region are displayed in Fig. 10. In Fig. 9, one can recognize two suspicious areas for which biopsies were taken, at 10 and 4 o'clock from the center of the os. Figure 11 shows a full image of the same patient with manual annotations by a doctor. Both biopsy locations are visible (small circles).

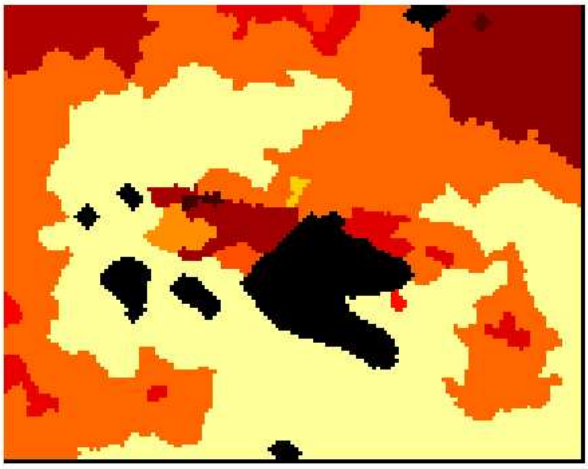

Fig. 8. Segmentation masks obtained with the clustering technique.

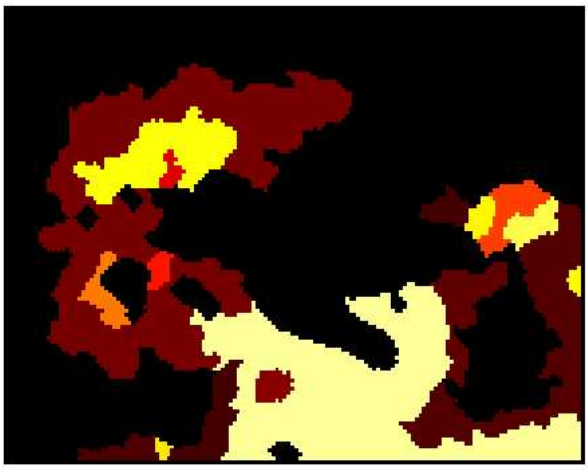

Fig. 9. Segmentation of background lesion with the clustering technique.

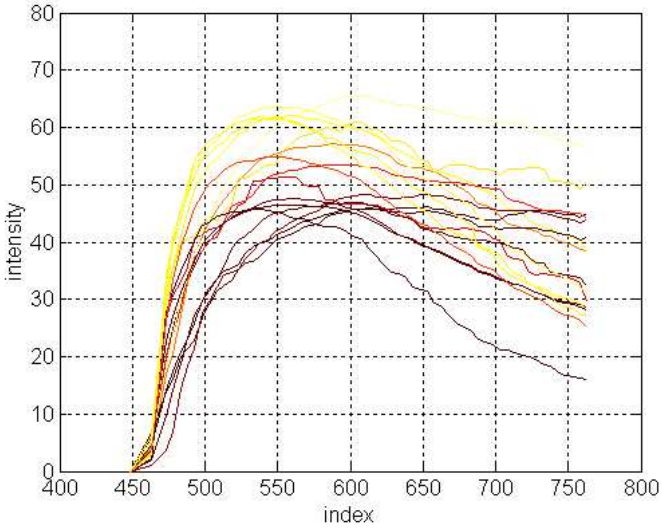

Fig. 10. Mean whiteness curves corresponding to the masks displayed in Fig. 9.

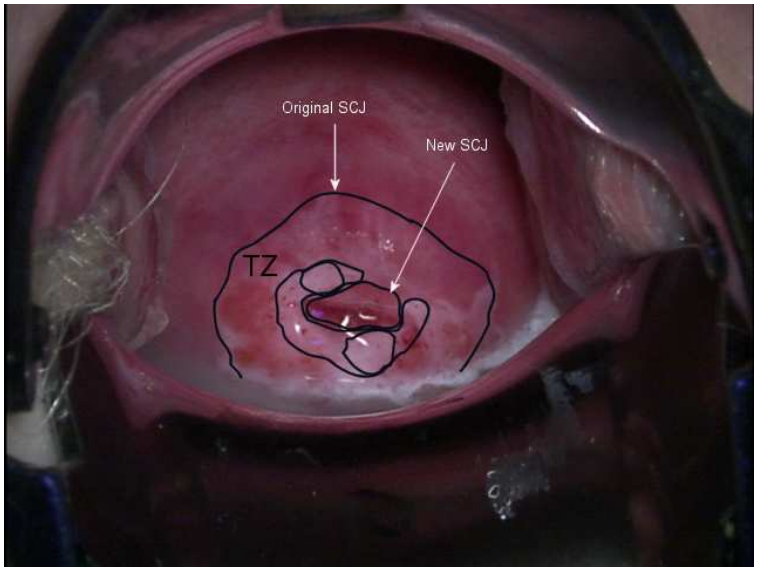

Fig. 11. Manual annotation of a patient.

VI. Data fitting and feature extraction

A. Model fitting

The whiteness signals extracted from the aligned sequences, even when they are the result of averaging many signals over a region, are rarely smooth. Before we extract features from the signals for the classification into different tissue classes, additional smoothness can be obtained by fitting a model to the data. Because no physiological model was available, we decided to use functions whose shape was close to that of a whiteness signal with a sharp uptake and a slow decay phase. Among others, we evaluated the use of a polynomial, a skewed Gaussian, and the sum of two tangent hyperbolic. Nonlinear least-squares techniques were used to fit these functions to the data [21]. Unfortunately, none of them performed well in all cases, and the number of parameters of some models was just too high. For the polynomial fit, it was difficult to decide which order was adequate for a given whiteness signal, leading most of the time to significant oscillations in the decay phase.

We developed a custom polynomial fitting technique where three polynomials of different order are fitted separately to the uptake, transition, and decay phases of the whitening signals. These phases are detected first, and once their bounds are known, a technique called *smooth phase fitting* is applied. Constraints are put at the phase boundaries in order to keep continuous first and second derivatives. The starting and ending slope are specified as well. The former is set to zero, since just before the acetic acid is applied, there is no whitening. For the end slope, a line is fitted to the last data samples. The different equations necessary to compute the smooth phase fitting are derived in the Appendix. Note that this technique can handle cases where the decay phase is missing and cases where both the transition and decay phases are missing.

In the following section, the technique used to split the data into three phases is explained.

B. Phase splitting

From now on we assume that mean whiteness signals are computed over candidate lesions outlined by the segmentation. Those mean signals are noted $\mu_{\mathbf{w}}$. In order to split the data into up to three phases, we need to detect the maximum whiteness location $n_{\max}$. A polynomial was fit to the whiteness signal. The maximum whiteness time was then estimated from this polynomial. Choosing the order of the polynomial is not straightforward, since it must be as

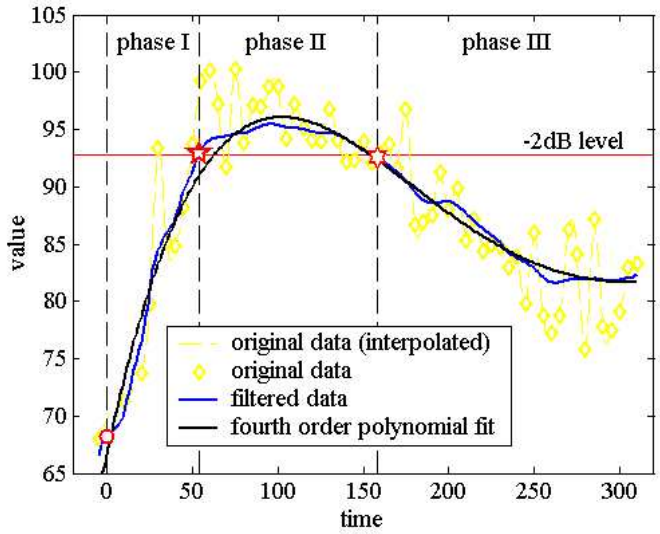

Fig. 12. Phase splitting.

small as possible to simplify the analysis while allowing for the "best" possible approximation of the data. In our experiments we decided to use polynomials of degree 4. This choice was done empirically and might be modified in the future. The polynomial coefficients are approximated in the least mean-square sense [22].

Four bounds $\{n_{p_0}, n_{p_1}, n_{p_2}, n_{p_3}\}$ are needed for the three phases, with $n_{p_0} < n_{p_1} \leq n_{p_2} \leq n_{p_3}$. Initially, the n_{p_1} and n_{p_2} bounds were chosen to be the points at -3 dB (86 %) of the signal amplitude measured at $n_{\max}$, where the amplitude is measured between $\mu_{\mathrm{w}}(n_1)$ and $\mu_{\mathrm{w}}(n_{\max})$. Note that the formula used to compute a ρ dB amplitude is given by:

$$\mu_\rho = \mu_{\mathbf{w}}(n_1) + \exp\left(\frac{\rho}{20}\right)\left[\mu_{\mathbf{w}}(n_{\max}) - \mu_{\mathbf{w}}(n_1)\right] \tag{19}$$

Since we work with discrete signals, there is usually no time step n for which $\mu_{\mathbf{w}}(n) = \mu_\rho$. Therefore, the closest index is used. While for the last bound we simply set $p_3 = N$, the lowest bound n_{p_0} is not chosen to be n_1 . The time-zero point is the time at -60 dB (5 %) amplitude. This decision was made because it happens very often that the first data sample correspond to images taken before the acetic acid was applied. By using the -60 dB amplitude point, this part of the data is automatically skipped.

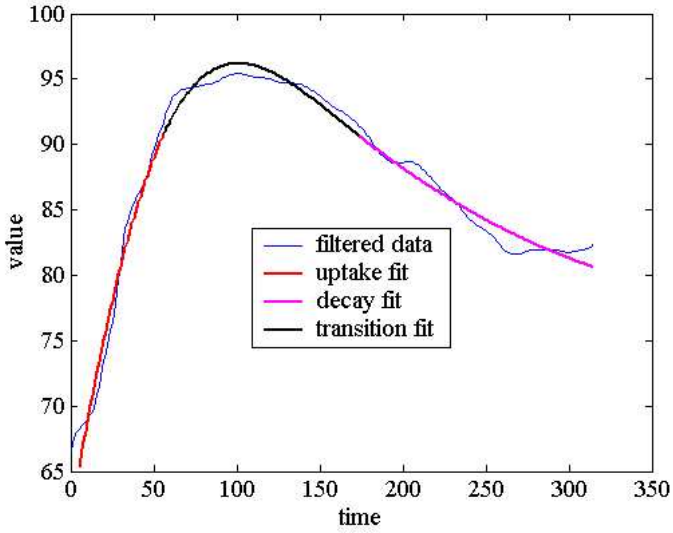

Fig. 13. Example of smooth phase fitting without start and end slope constraints.

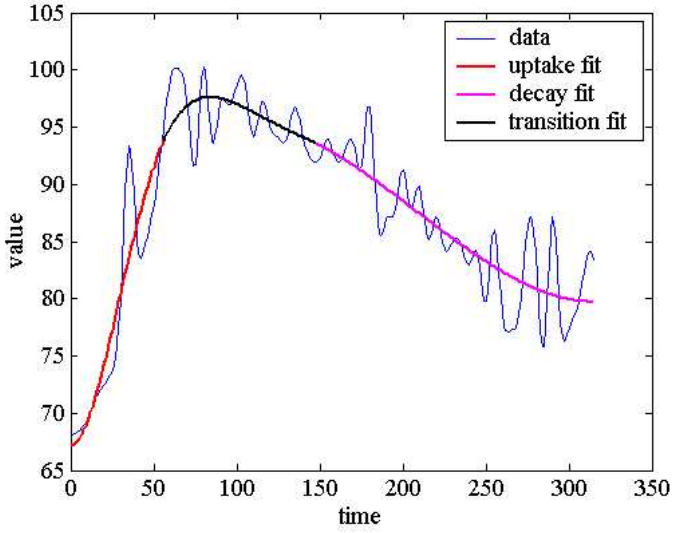

Fig. 14. Example of smooth phase fitting.

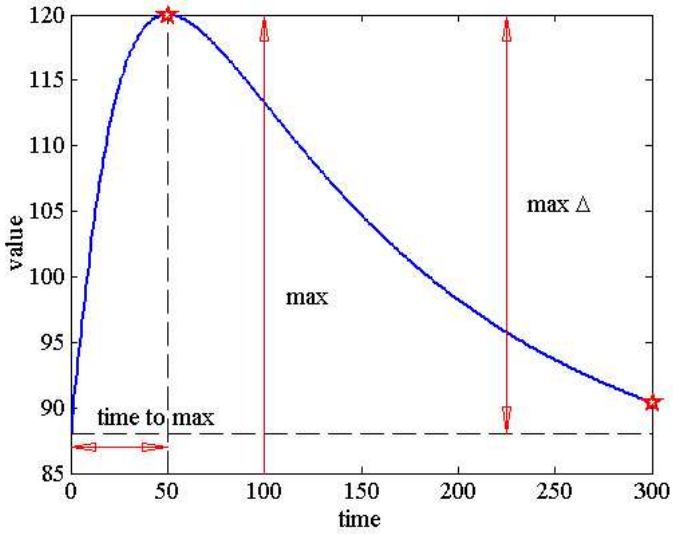

Fig. 15. Extracted features.

A few typical examples of smooth phase fitting are shown in Figs. 13 and 14. Note that in the case of Fig. 13 no constraint was set on the start and end slopes, i.e. the derivatives at the boundaries.

C. acetowhitening features

A number of features are computed from the fitted data. Those features are supposed to describe the shape of the whiteness curves as accurately as possible. The maximum whiteness point is of course an important point, and three features are computed to describe it: the time to the maximum whiteness point, its amplitude, and the corresponding delta amplitude, obtained by subtracting the time-zero whiteness value to the maximum amplitude. Those features are shown in Fig. 15. The slope and integral at particular time steps are used as features as well. The time steps are chosen to be the time at which the whiteness reaches a fraction of the delta amplitude. The different fractions are expressed in dBs. The extracted features are shown in Figs. 16 to 18. In our experiments, we extracted the slopes and integrals from -25 dB to -3 dB, for the uptake and decay phases. In addition to that, the raw data was also added to the feature vector.

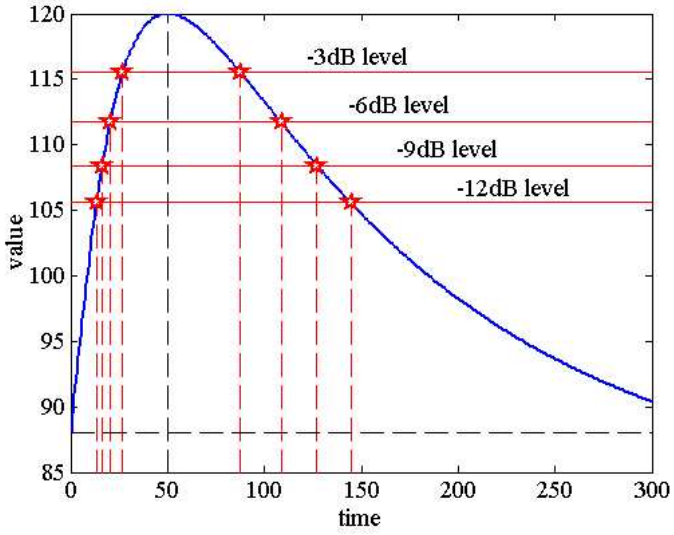

Fig. 16. dB points.

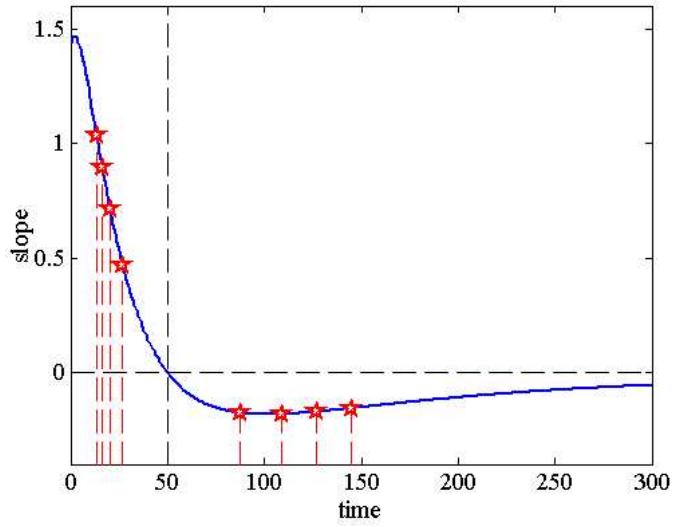

Fig. 17. Features from derivative of curve.

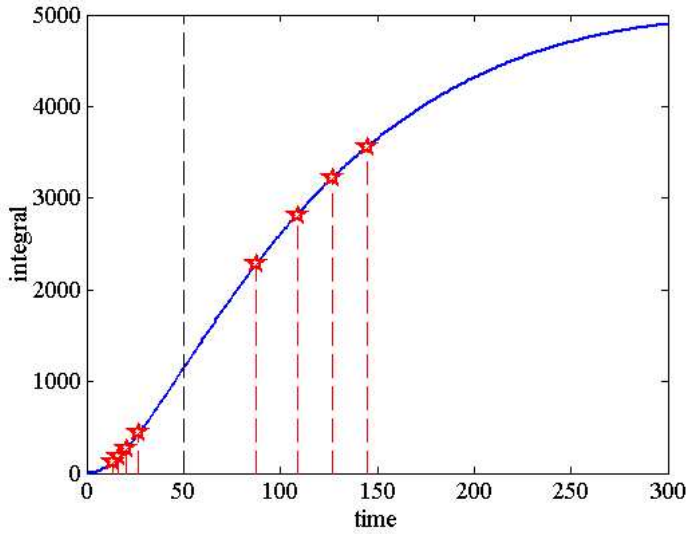

Fig. 18. Features from integral of curve.

VII. CLASSIFIER

The choice of an appropriate classifier is a very sensitive part in the design of computer-aided learning devices. Even though outstanding results were reported in machine learning applications using *artificial neural networks* (ANN) [23], the training and testing of such classifiers require large data sets. Because we did not have suficient patient data, especially high-grade cases, we employed a linear classification scheme. The design of more sophisticated classifiers is left for future investigations when more data will be available.

Identifying optimal parameters for the separation of classes was pursued by the application of a *discrimination function* DF to the data set. To evaluate the discriminating power of each feature, we defined the following *discrimination function*:

$$DF\left(f_k\right) = \frac{\left| \frac{\sum\limits_{i \in \mathbb{S}_1} f_{k,i}}{n_1} - \frac{\sum\limits_{j \in \mathbb{S}_2} f_{k,j}}{n_2} \right|}{\sqrt{\frac{\sigma_{k,1}^2 + \sigma_{k,2}^2}{n_1 + n_2}}} \tag{20}$$

where $\mathbb{S}_1$ and $\mathbb{S}_2$ are the two subsets to be discerned using feature f_k, σ_1 and σ_1 are the standard deviations of f_k in $\mathbb{S}_1$ and $\mathbb{S}_2$, respectively, and n_1 and n_2 are the number of samples in $\mathbb{S}_1$ and

$\mathbb{S}_2$, respectively. The numerator measures the difference in the mean f_k. The denominator is the probable error of the sum of the feature standard deviations. A high difference between the feature of interest and a small error in the standard deviation, would maximize the discrimination between class $\mathbb{S}_1$ and class $\mathbb{S}_2$. As a result, values greater than one provide increasing discrimination potential, while values less than one are dominated by the mean standard deviation error.

VIII. RESULTS

In this study, 244 patients were enrolled of which 44 presented with high-grade lesions, i.e CIN II-III lesions. In this set of patients, 21 cases were deemed *unusable* because of procedural difficulties: for example, biopsies were obscured by the specula and hence biopsy locations were not recorded on the video, cervical images were out of focus, or the disease was located in the cervical canal and hence was not observed in the collected images. For purposes of analysis the following data was used:

- 26 *high grade* pathology-confirmed lesions (CIN II-III) from 22 unique patients,
- 16 *low grade* pathology-confirmed lesions (CIN I) from 13 unique patients,
- 11 lesions evaluated by pathology as *no evidence of disease* (NED) from 10 unique patients (false positive colposcopies),
- 5 lesions evaluated by pathology as *metaplasia* (MET) from 5 unique patients.

The average polynomial fit kinetic curves of each pathology confirmed tissue class are displayed in Fig. 19. The green areas show the minimum and maximum values obtained at each time step, and the yellow areas are defined by the lower and upper quartiles [1] of the data at each time step. The continuous red lines correspond to the average data, and the dashed red curves to the standard deviations. Though there are distinctive trends produced by the mean curves in each class (CIN II-III, CIN I, NED and MET) the upper and lower bound curves show that the amplitude of the whitening alone was not sufficient for good separation of classes.

The discrimination function, for separating CIN II-III from the combination of CIN I, NED and MET, based upon the three phase fitted rising and decaying dB point slopes, is shown in Fig. 20. As seen in this figure, the region of the highest discrimination is calculated to occur between -12 and -17 dB on the decaying slope of the kinetic curve. Similarly, the discrimination

[1] 25 % and 75 % values of the ordered data, respectively

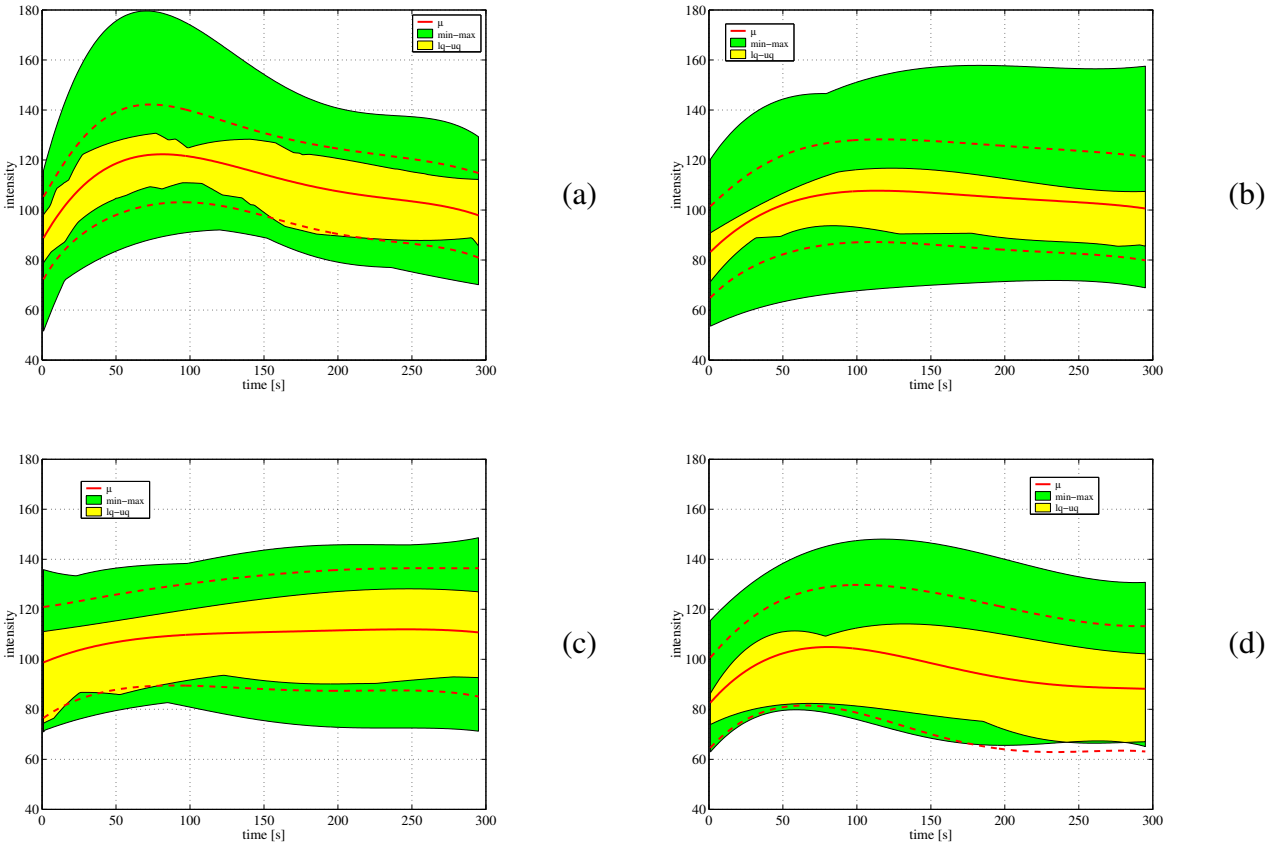

Fig. 19. Average kinetic curves for (a) high grade, (b) low grade, (c) NED, and (d) metaplasia lesions.

function, for separating CIN II-III from the combination of CIN I, NED and MET, based upon the derivative of the polynomial fitted acetowhitening kinetic curves, is shown in Fig. 22. As seen in this figure, the time of the highest discrimination is calculated to occur at 195 seconds.

In both discrimination functions it is noted that the functions are not sharply oscillatory and that the maximums span a wide temporal region, suggesting that the maximum discrimination regions are not a result of a statistical coincidence or noise, but rather a meaningful result of the acetowhitening kinetics curves. The results of using dB points and polynomial fitted acetowhitening curve time points as a classifying feature in the statistical analysis program SyStat [24] are shown in Figs. 21 and 23, respectively. The resulting accuracies, calculated using a linear regression multivariate analysis [25], track well with the result of the DF. Such

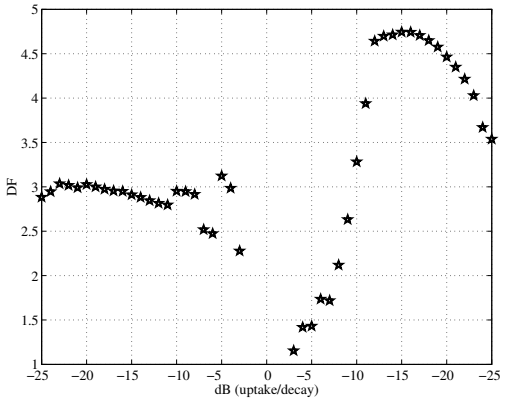

Fig. 20. Discrimination function for separating CIN II-III from CIN I, NED, and MET, based upon the dB slopes features.

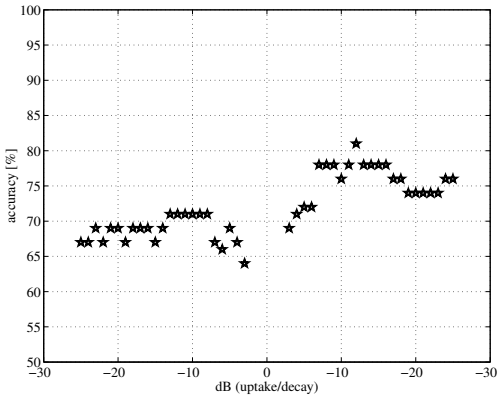

Fig. 21. Same as Fig 20 with the statistical analysis program SyStat.

Actual \ Predicted	CIN II-III	CIN I/NED/MET	Accuracy [%]
CIN II-III	17	9	65
CIN I/NED/MET	4	28	88

TABLE I

RESULT OF LDA USING A SINGLE FEATURE, THE DECAYING SLOPE AT -16 DB.

a correlation to the DF further suggests that the maximum DF regions may be useful for discrimination between CIN II-III and the combination of CIN I, NED, and MET.

Discrimination functions were calculated for all of the features, including the intensity, slope, time, and integral related features, as well as the slope at each time point of the polynomial fitted data (see Section VI). In addition metrics such as normalized fitted data, derivatives of the normalized data and integrals of the normalized data were evaluated using the DF. The formula used to normalize the fitted data was:

$$\frac{\mathrm{p}_{\max} - \mathrm{p}\left(n_{p_0}\right)}{\mathrm{p}\left(n_{p_0}\right)}\ , \tag{21}$$

where p is the phase fitted polynomial and $\mathrm{p}_{\max}$ is the maximum value reached inside the time bounds.

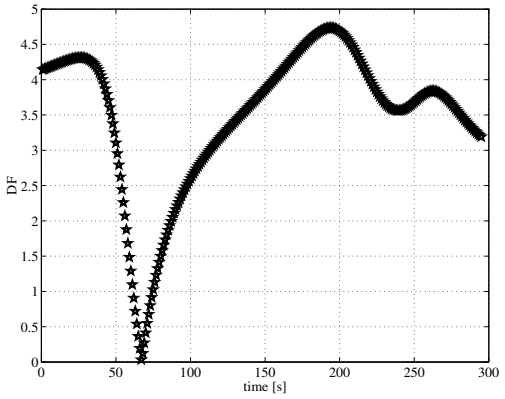

Fig. 22. Discrimination function for separating CIN II-III from CIN I, NED, and MET, based upon the time slopes features.

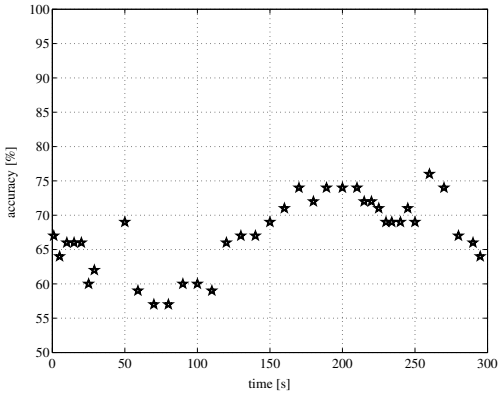

Fig. 23. Same as Fig 22 with the statistical analysis program SyStat.

Actual \ Predicted	CIN II-III	CIN I/NED/MET	Accuracy [%]
CIN II-III	16	10	62
CIN I/NED/MET	8	24	75

TABLE II

RESULT OF LDA USING A SINGLE FEATURE, THE SLOPE OF THE FITTED POLYNOMIAL AT 235 SECONDS.

The *decaying* dB slopes and the slopes of the polynomial fitted data, as shown in Figs. 20 and 22 respectively, displayed the highest DF values and therefore should prove the most useful in a linear classifier. The -11 to -18 dB slopes were imported into the linear discriminant analysis software. A complete linear discriminant analysis was performed on each slope individually. The range from -11 dB to -16 dB showed the highest accuracy for discriminating CIN II-III from CIN I, NED and MET. Other than the -12 dB point, which gave an accuracy of 81 %, each point in this range gave an accuracy of 78 %. These points are shown in Fig 21. An example of the LDA output is shown in Table 1 for the decaying -16 dB slope.

A complete linear discriminant analysis was also performed, at 5 to 10 second intervals, on the derivative of the polynomial fitted data surrounding the DF maximum. The range from 150 to 295 seconds after the application of acetic acid showed a mean accuracy for discriminating

Actual \ Predicted	CIN II-III	CIN I/NED/MET	Accuracy [%]
CIN II-III	23	3	88
CIN I/NED/MET	4	28	88

TABLE III

RESULT OF BACKWARD STEP LDA USING THE DECAYING SLOPE AT -16 DB AND THE SLOPE OF THE FITTED POLYNOMIAL AT 235 SECONDS.

CIN II-III from CIN I, NED and MET of 71 % $\pm$ 3 %. These points are shown in Fig. 22. Examples of the LDA output are shown in Tables I and II for the slopes at the -16 dB points and at 235 seconds after the application of acetic acid, respectively.

Improving classification with the use of two parameters was explored. The results of using the slopes at the decaying -16 dB point and the polynomial fitted 235 second concurrently, in a backward step LDA [25], improved the accuracy of separation to 88 %. Results are shown in Table III.

IX. DISCUSSION & CONCLUSIONS

This study was designed to measure multiple parameters of the acetowhitening process to determine if they can be used to improve the accurate classification of high-grade lesions and thus increase the sensitivity and specificity of traditional colposcopy.

The acetowhitening kinetic curves shown in Fig. 19 represent the average tissue specific curves for the data collected in this study. This initial analysis of the kinetic data concluded that a simple tissue whitening intensity was not sufficient to discriminate between tissue classes. However, the source of variability in the kinetic curves is of interest. Patient-to-patient intensity fluctuations may be a result of experimental conditions. For example, changes in the lighting conditions of the cervix associated with colposcope placement, shadowing caused by surrounding vaginal walls, and speculum placement may contribute to the variations observed in the collected data. Since, in the current implementation of the instrument, we do not have a measure of the absolute light intensity at the cervix, it is not possible to correct for fluctuations in lighting conditions. To compensate a normalization was performed on each individual kinetic curve using Eq. 21. This normalization evaluates the relative change in the intensity in relation to the pre-acetic acid

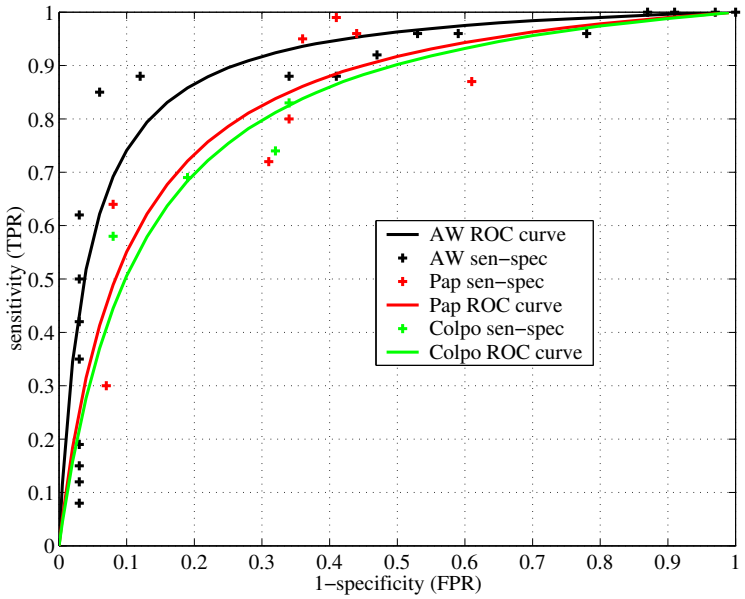

Fig. 24. ROC curve of acetowhitening study vs. pap/colposcopy.

intensity. While this normalization procedure helped to improve separation of the tissue classes, it was not sufficient to exclude overlap of the standard deviations. This suggests that the patient-to-patient variation in intensity may arise from variables other than experimental conditions. Variables that are intrinsic to the patient, such as age, race, or the stage of menstruation, may also contribute to the variability of the data [26], [27]. Such factors will be studied in the future.

Given the error associated with the features using the intensity of the acetowhitening kinetic curves, discrimination functions were calculated for other potential features extracted from the three-phase fit curves and the fourth-order polynomial fit curves. The highest values of the DF of all the investigated features were associated with the decaying dB point's slope and the derivative of the polynomial fit kinetic curves, shown in Figs. 20 to 23. From theses curves, all combinations of the features surrounding the peaks of the corresponding DF were entered as parameters in a jackknifed LDA algorithm. The combination of the decaying -16 dB slope and the slope of the polynomial fit at 235 seconds, produced the highest classification accuracy, with a sensitivity and a specificity of 88 %. The classification algorithm used the discriminant line $y = -0.9282x - 0.1348$ to produce the above accuracy of 88 %, where the y-axis corresponds to

the decaying -16 dB slope feature and the x-axis to the 235 seconds slope feature. Moving this line through the two-dimensional data space, while maintaining the slope of the line, produces all possible combinations of specificities and sensitivities. These are plotted as the true positive ratio (TPR) and the false positive ratio (FPR) [28] in Fig. 24.

This same data was used to calculate the *receiver operator characteristic* (ROC) curve [28], shown in black in Fig 24. The ROC curve represents a fit to the accuracies calculated using the decaying -16 dB slope and the polynomial fitted slope at 235 seconds (black $+$). This curve is calculated using the method described in [28]. The Q point for this curve occurs at 84 %. It is the point where sensitivity equals specificity. For comparison, Fig. 24 also shows the representative experimental ROC curves for colposcopy [1] and Pap smear [29] (green and red respectively). The Q point for colposcopy and Pap, calculated in a similar manner as above, is 75 % and 77 % respectively. The results presented herein, using the quantitatively measured kinetic acetowhitening curves, provide an overall increase in performance for discrimination of CIN II-III form CIN I, NED and MET, relative to colposcopy and Pap.

In conclusion, we have demonstrated that the quantitative measurement of the kinetic acetowhitening process has distinct advantages for the separation of CIN II-III from the combination of CIN I, NED, and MET. We have shown that the slopes, of the acetowhitening kinetic curve, at the decaying -16 dB point and the polynomial fitted 235 second point provides the highest tissue discrimination. By using slopes extracted at these two time points as the primary features in a multivariate analysis, we have shown a sensitivity of 88 % and a specificity of 88 %, resulting in an accuracy of 88 %. This demonstrates an improved specificity over colposcopy (specificity of 69 %) while almost maintaining sensitivity (sensitivity of 85 %) [1].

This encouraging result warrants further investigations into the quantitative measurement of the kinetic acetowhitening process. For example, the separation of diseased tissue may be improved by reducing the variability in the data. This variability may result from factors such as patient age [26] patient race, the stage of menstruation (i.e. estrogen levels), or more simply from inconsistent experimental conditions. Improved protocol design, including a better, more uniform method of acetic acid application, improved viewing of the cervix, normalized lighting conditions, accurate image registration, focus, and improved image quality may facilitate reduced data variability and improved diagnostic statistical accuracy. As an example, a fiduciary mark applied to the cervix prior to acid application could be used to account for fluctuating lighting

conditions, image registration, and focus.

In addition to reducing the variability between patients due to intrinsic and extrinsic factors, other features may be incorporated into the analysis that would improve diagnosis [30]. Structural tissue features, often noted by colposcopists during examinations, are used to aid in the location of suspicious regions. For example, the presence of *cuffed glands* is believed to be associated with a higher prevalence of disease [31]. Similarly, the variation in the whitening and the boarders within a suspicious region as well as the proximity to the new *squamous-columnar junction* (SCJ) [3] help to classify a region of interest. Some preliminary work, using image segmentation has shown encouraging results for computer-aided detection of structural features. Future automation of one or several of these important features may facilitate the accuracy of diagnosis.

Finally, other tissue types must be investigated. To date we have included CIN, NED and mixed MET. It is important that columnar tissue, delineated metaplastic tissue (such as mature and immature), invasive carcinoma, cysts, polyps, cervicitis, inflammation and other atypias [3] be included in our investigation. Larger enhanced clinical studies with the aforementioned potential improved diagnostics will be well suited for this type of intensive study.

Acknowledgements

We gratefully acknowledge the support of Ms. Jean Reichert whose clinical coordination was invaluable to this study. We also would like to thank the support staff of the Department of Obstetrics and Gynecology at the Dartmouth-Hitchcock Medical Center in Lebanon, New Hampshire, whose help was also crucial for the completion of this study. This work was supported by MediSpectra, Inc. (Lexington, MA) and the National Cancer Institute 1R44CA91618-01.

APPENDIX: SMOOTH PHASE FITTING

Let the orders of the uptake, transition, and decay polynomials be $q_1 \geq 2$, $q_2 \geq 2$, and $q_3 \geq 2$, respectively. The fitted function is then given by:

$$\mathrm{p}\left(n\right) = \begin{cases} \sum\limits_{k=0}^{q_1} a_k n^k & \text{if } n \in [n_{p_0}, n_{p_1}] \\ \sum\limits_{k=0}^{q_2} b_k n^k & \text{if } n \in [n_{p_1}, n_{p_2}] \\ \sum\limits_{k=0}^{q_3} c_k n^k & \text{if } n \in [n_{p_2}, n_{p_3}] \end{cases} . \tag{22}$$

Given a set of points

$$\mathbf{y} = \begin{pmatrix} \mu_{\mathbf{w}}\left(n_{p_0}\right) \\ \vdots \\ \mu_{\mathbf{w}}\left(n_{p_3}\right) \end{pmatrix} , \tag{23}$$

The system of linear equations is then given by:

$$\mathbf{A}\mathbf{x} = \mathbf{y} , \tag{24}$$

where

$$\mathbf{A} = \begin{pmatrix} n_{p_0}^0 & \cdots & n_{p_0}^{q_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ n_{p_1-1}^0 & \cdots & n_{p_1-1}^{q_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & n_{p_1}^0 & \cdots & n_{p_1}^{q_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & n_{p_2}^0 & \cdots & n_{p_2}^{q_2} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & n_{p_2+1}^0 & \cdots & n_{p_2+1}^{q_3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & n_{p_3}^0 & \cdots & n_{p_3}^{q_3} \end{pmatrix} \tag{25}$$

and

$$\mathbf{x} = \begin{pmatrix} a_0 & \cdots & a_{q_1} & b_0 & \cdots & b_{q_2} & c_0 & \cdots & c_{q_3} \end{pmatrix}^T . \tag{26}$$

Eq. 24 can be solved in the least mean-square sense for the vector of coefficients $\mathbf{x}$:

$$\mathbf{x} \approx \left(\mathbf{A}^T\mathbf{A}\right)\mathbf{A}^T\mathbf{y}. \tag{27}$$

In order to obtain a continuous function, as well as continuous first and second order derivatives at the phase boundaries, the following constraints are necessary:

$$\sum_{k=0}^{q_1} a_k n_{p_1}^k = \sum_{k=0}^{q_2} b_k n_{p_1}^k \tag{28}$$

$$\sum_{k=1}^{q_1} k a_k n_{p_1}^{k-1} = \sum_{k=1}^{q_2} k b_k n_{p_1}^{k-1} \tag{29}$$

$$\sum_{k=2}^{q_1} k\,(k-1)\,a_k n_{p_1}^{k-2} = \sum_{k=2}^{q_2} k(k-1) b_k n_{p_1}^{k-2} \tag{30}$$

$$\sum_{k=0}^{q_3} c_k n_{p_2}^k = \sum_{k=0}^{q_2} b_k n_{p_2}^k \tag{31}$$

$$\sum_{k=1}^{q_3} k c_k n_{p_2}^{k-1} = \sum_{k=1}^{q_2} k b_k n_{p_2}^{k-1} \tag{32}$$

$$\sum_{k=2}^{q_3} k\,(k-1)\,c_k n_{p_2}^{k-2} = \sum_{k=2}^{q_2} k(k-1) b_k n_{p_2}^{k-2}\ . \tag{33}$$

To set the slope at the signal boundaries, two additional constraints are necessary:

$$\sum_{k=1}^{q_1} k a_k n_{p_0}^{k-1} = d_0 \tag{34}$$

$$\sum_{k=1}^{q_3} k c_k n_{p_3}^{k-1} = d_3\ . \tag{35}$$

The set of constraints can then be written:

$$a_0 + a_1 n_{p_1} + a_2 n_{p_1}^2 - b_1 n_{p_1} - b_2 n_{p_1}^2 = -\sum_{k=3}^{q_1} a_k n_{p_1}^k + b_0 + \sum_{k=3}^{q_2} b_k n_{p_1}^k \quad (36)$$

$$a_1 + 2a_2 n_{p_1} - b_1 - 2b_2 n_{p_1} = -\sum_{k=3}^{q_1} k a_k n_{p_1}^{k-1} + \sum_{k=3}^{q_2} k b_k n_{p_1}^{k-1} \quad (37)$$

$$a_1 + 2a_2 n_{p_0} = -\sum_{k=3}^{q_1} k a_k n_{p_0}^{k-1} + d_0 \quad (38)$$

$$2a_2 - 2b_2 = -\sum_{k=3}^{q_1} k\,(k-1)\, a_k n_{p_1}^{k-2} + \sum_{k=3}^{q_2} k(k-1) b_k n_{p_1}^{k-2} \quad (39)$$

$$-b_1 n_{p_2} - b_2 n_{p_2}^2 + c_0 + c_1 n_{p_2} + c_2 n_{p_2}^2 = b_0 + \sum_{k=3}^{q_2} b_k n_{p_2}^k - \sum_{k=3}^{q_3} c_k n_{p_2}^k \quad (40)$$

$$-b_1 - 2b_2 n_{p_2} + c_1 + 2c_2 n_{p_2} = \sum_{k=3}^{q_2} k b_k n_{p_2}^{k-1} - \sum_{k=3}^{q_3} k c_k n_{p_2}^{k-1} \quad (41)$$

$$c_1 + 2c_2 n_{p_3} = -\sum_{k=3}^{q_3} k c_k n_{p_3}^{k-1} + d_3 \quad (42)$$

$$-2b_2 + 2c_2 = \sum_{k=3}^{q_2} k(k-1) b_k n_{p_2}^{k-2} - \sum_{k=3}^{q_3} k\,(k-1)\, c_k n_{p_2}^{k-2} \quad (43)$$

The matrix formulation is:

$$\bar{\mathbf{x}} = \mathbf{C}\mathbf{S}\tilde{\mathbf{x}} + \mathbf{C}\mathbf{d}\ , \quad (44)$$

where

$$\bar{\mathbf{x}} = \begin{pmatrix} a_0 & a_1 & a_2 & b_1 & b_2 & c_0 & c_1 & c_2 \end{pmatrix}^T \quad (45)$$

$$\tilde{\mathbf{x}} = \begin{pmatrix} a_3 & \cdots & a_{q_1} & b_0 & b_3 & \cdots & b_{q_2} & c_3 & \cdots & c_{q_3} \end{pmatrix}^T \quad (46)$$

$$\mathbf{d} = \begin{pmatrix} 0 & 0 & d_0 & 0 & 0 & 0 & d_3 & 0 \end{pmatrix}^T \quad (47)$$

and

$$\mathbf{C} = \begin{pmatrix} 1 & n_{p_1} & n_{p_1}^2 & -n_{p_1} & -n_{p_1}^2 & 0 & 0 & 0 \\ 0 & 1 & 2n_{p_1} & -1 & -2n_{p_1} & 0 & 0 & 0 \\ 0 & 1 & 2n_{p_0} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -n_{p_2} & -n_{p_2}^2 & 1 & n_{p_2} & n_{p_2}^2 \\ 0 & 0 & 0 & -1 & -2n_{p_2} & 0 & 1 & 2n_{p_2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2n_{p_3} \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 & 2 \end{pmatrix}^{-1} \tag{48}$$

$$\mathbf{S} = \begin{pmatrix} -n_{p_1}^3 & \cdots & -n_{p_1}^{q_1} & 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} & 0 & \cdots & 0 \\ -3n_{p_1}^2 & \cdots & -q_1 n_{p_1}^{q_1-1} & 0 & 3n_{p_1}^2 & \cdots & q_2 n_{p_1}^{q_2-1} & 0 & \cdots & 0 \\ -3n_{p_0}^2 & \cdots & -q_1 n_{p_0}^{q_1-1} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ -6n_{p_1} & \cdots & -q_1 (q_1-1) n_{p_1}^{q_1-2} & 0 & 6n_{p_1} & \cdots & q_2 (q_2-1) n_{p_1}^{q_2-2} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & n_{p_2}^3 & \cdots & n_{p_2}^{q_2} & -n_{p_2}^3 & \cdots & -n_{p_2}^{q_3} \\ 0 & \cdots & 0 & 0 & 3n_{p_2}^2 & \cdots & q_2 n_{p_2}^{q_2-1} & -3n_{p_2}^2 & \cdots & -q_3 n_{p_2}^{q_3-1} \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & -3n_{p_3}^2 & \cdots & -q_3 n_{p_3}^{q_3-1} \\ 0 & \cdots & 0 & 0 & 6n_{p_2} & \cdots & q_2 (q_2-1) n_{p_2}^{q_2-2} & -6n_{p_2} & \cdots & -q_3 (q_3-1) n_{p_2}^{q_3-2} \end{pmatrix} \tag{49}$$

The vector $\bar{\mathbf{x}}$ of constrained parameters can be computed with Eq. 44. Equation 24 can be rewritten as:

$$\mathbf{A}_1 \tilde{\mathbf{x}} + \mathbf{A}_2 \bar{\mathbf{x}} = \mathbf{y} \ , \tag{50}$$

where

$$\mathbf{A}_1 = \begin{pmatrix} n_{p_0}^3 & \cdots & n_{p_0}^{q_1} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ n_{p_1-1}^3 & \cdots & n_{p_1-1}^{q_1} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & n_{p_2}^3 & \cdots & n_{p_2}^{q_2} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & n_{p_2-1}^3 & \cdots & n_{p_2-1}^{q_3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & n_{p_3}^3 & \cdots & n_{p_3}^{q_3} \end{pmatrix} \tag{51}$$

$$\mathbf{A}_2 = \begin{pmatrix} 1 & n_{p_0} & n_{p_0}^2 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n_{p_1-1} & n_{p_1-1}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & n_{p_1} & n_{p_1}^2 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & n_{p_2} & n_{p_2}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & n_{p_2+1} & n_{p_2+1}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 & n_{p_3} & n_{p_3}^2 \end{pmatrix} \tag{52}$$

Let us define $\tilde{\mathbf{y}}$ as:

$$\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{A}_2\mathbf{C}\mathbf{d} \ . \tag{53}$$

Equation 50 becomes:

$$\left(\mathbf{A}_1 + \mathbf{A}_2\mathbf{C}\mathbf{S}\right)\tilde{\mathbf{x}} = \tilde{\mathbf{y}} \ . \tag{54}$$

Finally, the solution in the least mean-squares sense becomes:

$$\tilde{\mathbf{x}} \approx \left(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}\right)^{-1}\tilde{\mathbf{A}}^T\tilde{\mathbf{y}} \ , \tag{55}$$

with

$$\tilde{\mathbf{A}} = \mathbf{A}_1 + \mathbf{A}_2\mathbf{C}\mathbf{S} \ . \tag{56}$$

To summarize, the different processing steps of this constrained smooth phase fitting algorithm are the following:

1) Compute $\mathbf{C}$ with Eq. 48,
2) Compute $\mathbf{S}$ with Eq. 49,
3) Compute $\mathbf{A}_1$ with Eq. 51,
4) Compute $\mathbf{A}_2$ with Eq. 52,
5) Compute $\tilde{\mathbf{A}}$ with Eq. 56,
6) Compute $\mathbf{d}$ with Eq. 47,
7) Compute $\tilde{\mathbf{y}}$ with Eq. 53,
8) Compute the polynomial coefficients using Eqs. 55 and 44.

If there is no decay phase, then the constraints reduce to:

$$\sum_{k=0}^{q_1} a_k n_{p_1}^k = \sum_{k=0}^{q_2} b_k n_{p_1}^k \tag{57}$$

$$\sum_{k=1}^{q_1} k a_k n_{p_1}^{k-1} = \sum_{k=1}^{q_2} k b_k n_{p_1}^{k-1} \tag{58}$$

$$\sum_{k=1}^{q_1} k a_k n_{p_0}^{k-1} = d_0 \tag{59}$$

$$\sum_{k=2}^{q_1} k\,(k-1)\,a_k n_{p_1}^{k-2} = \sum_{k=2}^{q_2} k(k-1) b_k n_{p_1}^{k-2} \tag{60}$$

$$\sum_{k=1}^{q_3} k b_k n_{p_2}^{k-1} = d_2, \tag{61}$$

where n_{p_2} is set to be the last time sample. The matrix formulation is:

$$\begin{pmatrix} 1 & n_{p_1} & n_{p_1}^2 & -n_{p_1} & -n_{p_1}^2 \\ 0 & 1 & 2n_{p_1} & -1 & -2n_{p_1} \\ 0 & 1 & 2n_{p_0} & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2n_{p_2} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} =$$

$$\begin{pmatrix} -n_{p_1}^3 & \cdots & -n_{p_1}^{q_1} & 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} \\ -3n_{p_1}^2 & \cdots & -q_1 n_{p_1}^{q_1-1} & 0 & 3n_{p_1}^2 & \cdots & q_2 n_{p_1}^{q_2-1} \\ -3n_{p_0}^2 & \cdots & -q_1 n_{p_0}^{q_1-1} & 0 & 0 & \cdots & 0 \\ -6n_{p_1} & \cdots & -q_1 (q_1-1) n_{p_1}^{q_1-2} & 0 & 6n_{p_1} & \cdots & q_2 (q_2-1) n_{p_1}^{q_2-2} \\ 0 & \cdots & 0 & 0 & -3n_{p_2}^2 & \cdots & -q_2 n_{p_2}^{q_2-1} \end{pmatrix} \begin{pmatrix} a_3 \\ \vdots \\ a_{q_1} \\ b_0 \\ b_3 \\ \vdots \\ b_{q_2} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ d_0 \\ 0 \\ d_2 \end{pmatrix} . \tag{62}$$

Let the parameter vectors be:

$$\bar{\mathbf{x}} = \begin{pmatrix} a_0 & a_1 & a_2 & b_1 & b_2 \end{pmatrix}^T \tag{63}$$

$$\tilde{\mathbf{x}} = \begin{pmatrix} a_3 & \cdots & a_{q_1} & b_0 & b_3 & \cdots & b_{q_2} \end{pmatrix}^T . \tag{64}$$

Then, Eqs. 48, 49, 51, 52, and 47 become:

$$\mathbf{C} = \begin{pmatrix} 1 & n_{p_1} & n_{p_1}^2 & -n_{p_1} & -n_{p_1}^2 \\ 0 & 1 & 2n_{p_1} & -1 & -2n_{p_1} \\ 0 & 1 & 2n_{p_0} & 0 & 0 \\ 0 & 0 & 2 & 0 & -2 \\ 0 & 0 & 0 & 1 & 2n_{p_2} \end{pmatrix}^{-1} \tag{65}$$

$$\mathbf{S} = \begin{pmatrix} -n_{p_1}^3 & \cdots & -n_{p_1}^{q_1} & 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} \\ -3n_{p_1}^2 & \cdots & -q_1 n_{p_1}^{q_1-1} & 0 & 3n_{p_1}^2 & \cdots & q_2 n_{p_1}^{q_2-1} \\ -3n_{p_0}^2 & \cdots & -q_1 n_{p_0}^{q_1-1} & 0 & 0 & \cdots & 0 \\ -6n_{p_1} & \cdots & -q_1 (q_1-1) n_{p_1}^{q_1-2} & 0 & 6n_{p_1} & \cdots & q_2 (q_2-1) n_{p_1}^{q_2-2} \\ 0 & \cdots & 0 & 0 & -3n_{p_2}^2 & \cdots & -q_2 n_{p_2}^{q_2-1} \end{pmatrix} \tag{66}$$

$$\mathbf{A}_1 = \begin{pmatrix} n_{p_0}^3 & \cdots & n_{p_0}^{q_1} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ n_{p_1-1}^3 & \cdots & n_{p_1-1}^{q_1} & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & n_{p_2}^3 & \cdots & n_{p_2}^{q_2} \end{pmatrix} \tag{67}$$

$$\mathbf{A}_2 = \begin{pmatrix} 1 & n_{p_0} & n_{p_0}^2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & n_{p_1-1} & n_{p_1-1}^2 & 0 & 0 \\ 0 & 0 & 0 & n_{p_1} & n_{p_1}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & n_{p_2} & n_{p_2}^2 \end{pmatrix} \tag{68}$$

$$\mathbf{d} = \begin{pmatrix} 0 & 0 & d_0 & 0 & d_2 \end{pmatrix}^T , \tag{69}$$

respectively. If there is only an uptake phase, then the constraints are:

$$\sum_{k=1}^{q_1} k a_k n_{p_0}^{k-1} = d_0 \tag{70}$$

$$\sum_{k=1}^{q_1} k a_k n_{p_1}^{k-1} = d_1 , \tag{71}$$

where n_{p_1} is set to be the last time sample. The matrix formulation is:

$$\begin{pmatrix} 1 & 2n_{p_0} \\ 1 & 2n_{p_1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 & -3n_{p_0}^2 & \cdots & -q_1 n_{p_0}^{q_1-1} \\ 0 & -3n_{p_1}^2 & \cdots & -q_1 n_{p_1}^{q_1-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_3 \\ \vdots \\ a_{q_1} \end{pmatrix} + \begin{pmatrix} d_0 \\ d_1 \end{pmatrix} . \tag{72}$$

Let the parameter vectors be $\bar{\mathbf{x}} = \begin{pmatrix} a_1 & a_2 \end{pmatrix}^T$ and $\tilde{\mathbf{x}} = \begin{pmatrix} a_0 & a_3 & \cdots & a_{q_1} \end{pmatrix}^T$. Then, Eqs. 48, 49, 51, 52, and 47 become:

$$\mathbf{C} = \begin{pmatrix} 1 & 2n_{p_0} \\ 1 & 2n_{p_1} \end{pmatrix}^{-1} = \frac{1}{2\left(n_{p_1} - n_{p_0}\right)} \begin{pmatrix} 2n_{p_1} & -2n_{p_0} \\ -1 & 1 \end{pmatrix} \tag{73}$$

$$\mathbf{S} = \begin{pmatrix} 0 & -3n_{p_0}^2 & \cdots & -q_1 n_{p_0}^{q_1-1} \\ 0 & -3n_{p_1}^2 & \cdots & -q_1 n_{p_1}^{q_1-1} \end{pmatrix} \tag{74}$$

$$\mathbf{A}_1 = \begin{pmatrix} 1 & n_{p_0}^3 & \cdots & n_{p_0}^{q_1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & n_{p_1}^3 & \cdots & n_{p_1}^{q_2} \end{pmatrix} \tag{75}$$

$$\mathbf{A}_2 = \begin{pmatrix} n_{p_0} & n_{p_0}^2 \\ \vdots & \vdots \\ n_{p_1} & n_{p_1}^2 \end{pmatrix} \tag{76}$$

$$\mathbf{d} = \begin{pmatrix} d_0 & d_1 \end{pmatrix}^T \tag{77}$$

Note that in most cases the matrix $\mathbf{A}^T\mathbf{A}$ is close to singular. It is therefore suggested to use a singular value decomposition (SVD) to invert that matrix. The SVD does the following decomposition:

$$\mathbf{A}^T\mathbf{A} = \mathbf{U}\mathbf{W}\mathbf{V}^T \tag{78}$$

where $\mathbf{W}$ is a diagonal matrix containing the singular values. The inverse of $\mathbf{W}$ is then simply obtained by replacing the diagonal elements with their inverse. The matrices $\mathbf{U}$ and $\mathbf{V}^T$ are inverted numerically, and finally one can compute $\left(\mathbf{A}^T\mathbf{A}\right)^{-1} = \left(\mathbf{V}^T\right)^{-1}\mathbf{W}^{-1}\mathbf{U}^{-1}$.

All routines necessary for inverting matrices and for computing SVDs are available in [32].

REFERENCES

[1] M. F. Mitchell, D. Schottenfeld, G. Tortolero-Luna, S. Cantor, and R. Richards-Kortum, "Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis," *Obstet Gynecol*, vol. 91, no. 4, pp. 626–631, 1998.

[2] L. Burke, A. Dimoca, E. Orfanudaki, and E. Koumantakis, *Colposcopy: text and atlas*. Appleton and Lange, 1991.

[3] A. Singer and J. M. Monaghan, *Lower genital tract precancer*, 2nd ed. Blackwell Science, 2000.

[4] R. Drezek, A. Dunn, and R. Richards-Kortum, "Light scattering from cells: finite-difference time-domain simulations and goniometric measurements," *Applied Optics*, vol. 36, no. 16, pp. 3651–3661, 1999.

[5] M. Rajadhyaksha, G. Meaker, and S. Gonzalez, "Confocal microscopy of excised human skin using acetic acid and cross polarization: rapid detection of non-melonoma skin cancers," *SPIE Proceedings*, vol. 3907, pp. 84–88, 2000.

[6] E. Hopman, F. Voorhorst, P. Kenemans, C. Meyers, and T. Helmerhorst, "Observer agreement on interpreting colposcopic images of cin," *Gynecologic Oncology*, vol. 58, pp. 206–209, 1995.

[7] L. S. Massad and Y. C. Collins, "Strength of correlations between colposcopic impression and biopsy histology," *Gynecologic Oncology*, pp. 424–428, 2003.

[8] C. Balas, G. C. Themelis, E. P. Prokopakis, I. Orfanudaki, E. Koumantakis, and E. S. Helidonis, "In vivo detection and staging of epithelial dysplasias and malignancies based on the quantitative assessment of acetic acid-tissue interaction kinetics," *Journal of Photochemistry and Photobiology*, pp. 1–5, 1999.

[9] C. Balas, "A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix," *IEEE Transactions on biomedical engineering*, vol. 48, no. 1, pp. 96–104, 2001.

[10] K. Jack, *Video demystified*, 3rd ed. LLH Technology, 2001.

[11] G. Wyszecki and W. S. Stiles, *Color Science: concepts and methods, quantitative data and formulae*, 2nd ed. Wiley Classics Library, 2000.

[12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Addison Wesley, 1992.

[13] L. Vincent, *Morphological Algorithms*. Marcel Dekker, Inc, 1993, ch. 8, pp. 255–288.

[14] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, July 1990.

[15] S. Beucher and F. Meyer, *The Morphological Approach to Segmentation: The Watershed Transformation*. Marcel Dekker, Inc, 1993, ch. 12, pp. 433–481.

[16] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Transactions on Patterns Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583–598, June 1991.

[17] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, 1981.

[18] P. Schmid, "Segmentation of digitized dermatoscopic images by 2d color clustering," *IEEE Transactions on Medical Imaging*, vol. 18, no. 2, pp. 164–171, 1999.

[19] S. Beucher, "Watershed, hierarchical segmentation and waterfall algorithm," in *Mathematical Morphology and Its Applications to Image Processing*. Kluwer Academics, 1994.

[20] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Transactions on Image Processing*, vol. 2, no. 2, pp. 176–201, 1993.

[21] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, ser. Classics in Applied Mathematics. SIAM, 1996.

[22] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.

[23] T. M. Mitchell, *Machine learning*. McGraw-Hill, 1997.

[24] *SyStat 8.0 for Windows*, Systat Software, Inc., Point Richmond, USA, www.systat.com.

[25] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[26] D. M. Zahm, I. Nindl, C. Greinke, H. Hoyer, and A. Schneider, "Colposcopic appearance of cervical intraepithelial neoplasia is age dependent," *Am J Obstet Gynecol*, vol. 179, no. 5, pp. 1298–1304, 1998.

[27] C. MacAulay, "Variations of fluorescence spectroscopy during the menstrual cycle," *Optics Express*, vol. 10, no. 12, 2002.

[28] B. Littenberg and L. E. Moses, "Estimating diagnostic accuracy from multiple conflicting reports:a new meta-analytic method," *Med Decis Making*, vol. 13, pp. 313–321, 1993.

[29] K. Nanda, D. C. McCrory, E. R. Myers, L. A. Bastian, V. Hasselblad, J. D. Hickey, and D. B. Matchar, "Accuracy of the papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review," *Ann Intern Med*, vol. 132, no. 10, pp. 810–819, 2000.

[30] R. Reid and P. Scalzi, "An imporved colposcopic index for differentiating benign papillomaviral infections from high-grade cervical intraepithelial neoplasia," *Am J Obstet Gynecol*, vol. 153, pp. 611–618, 1985.

[31] Y. Kishi, S. Inui, and Y. Sakamoto, "Colposcopic findings of gland openings in cervical carcinoma: Their histological background," *Int J Gynaecol Obstet*, vol. 25, pp. 223–233, 1987.

[32] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. Cambridge University Press, 1999.