

Automatic Image Quality Assessment for Uterine Cervical Imagery

Jia Gu and Wenjing Li. (Email: jgu@sti-hawaii.com)
STI Medical Systems, 733 Bishop Street, Suite 3100, Honolulu, HI 96813, USA

Abstract

Uterine cervical cancer is the second most common cancer among women worldwide. However, its death rate can be dramatically reduced by appropriate treatment, if early detection is available. We are developing a Computer-Aided-Diagnosis (CAD) system to facilitate colposcopic examinations for cervical cancer screening and diagnosis. Unfortunately, the effort to develop fully automated cervical cancer diagnostic algorithms is hindered by the paucity of high quality, standardized imaging data. The limited quality of cervical imagery can be attributed to several factors, including: incorrect instrumental settings or positioning, glint (specular reflection), blur due to poor focus, and physical contaminants. Glint eliminates the color information in affected pixels and can therefore introduce artifacts in feature extraction algorithms. Instrumental settings that result in an inadequate dynamic range or an overly constrained region of interest can reduce or eliminate pixel information and thus make image analysis algorithms unreliable. Poor focus causes image blur with a consequent loss of texture information. In addition, a variety of physical contaminants, such as blood, can obscure the desired scene and reduce or eliminate diagnostic information from affected areas. Thus, automated feedback should be provided to the colposcopist as a means to promote corrective actions. In this paper, we describe automated image quality assessment techniques, which include region of interest detection and assessment, contrast dynamic range assessment, blur detection, and contaminant detection. We have tested these algorithms using clinical colposcopic imagery, and plan to implement these algorithms in a CAD system designed to simplify high quality data acquisition. Moreover, these algorithms may also be suitable for image quality assessment in telemedicine applications.

Key words: *Automatic Image quality assessment, Cervical Cancer, Colposcopy, Telemedicine, Computer-Aided-Diagnosis*

1. INTRODUCTION

Uterine cervical cancer is the second most common cancer in women worldwide, with nearly 500,000 new cases and over 270,000 deaths annually¹. Colposcopy is a diagnostic method used to detect cancer precursors and cancer of the uterine cervix⁴. Computer-Aided-Diagnosis (CAD) for colposcopy represents a new application of medical image processing. STI Medical Systems is developing a CAD system that mimics the diagnostic process used by colposcopists to assess the severity of abnormalities⁵. Scoring schemes, like the Reid's colposcopic index³, aid in making colposcopic diagnoses based on various features, including acetowhitening, vessel patterns and lesion margins. These features are individually assessed and scored before the scores of all features are combined to yield a composite score that grades disease severity. However, the quality of the images must be assessed before further analysis, to ensure reliable diagnostic performance. Hence, we present a systematic framework of algorithms to automatically assess cervical images acquired from a digital colposcope. The filtered dataset can then be used for CAD algorithms. Moreover, we are extending this technique for application to cervical cancer diagnosis via telemedicine, where it can be used to control and improve image acquisition quality.

The limited quality of cervical imagery can be attributed to several factors, including: incorrect instrumental settings or positioning, glint (specular reflection), blur due to poor focus, and physical contaminants. Glint eliminates the color information in affected pixels and can therefore introduce artifacts in feature extraction algorithms. Instrumental settings that result in an inadequate dynamic range or overly constrained region of interest can reduce or eliminate pixel information and thus make image analysis algorithms unreliable. Poor focus causes image blur with a consequent loss of texture information. In addition, a variety of physical contaminants, such as blood can obscure the desired scene and reduce or

eliminate diagnostic information from affected areas. Thus, this paper describes series of image processing algorithms which includes image region assessment, image contrast assessment, image focus assessment and contamination detection. Note that our data is acquired from STI digital Colposcope¹⁶.

2. ALGORITHM DESCRIPTION

The image processing algorithm described in this paper starts from RGB (Red-Green-Blue color space) images from 149 human subjects. The input image for the algorithm is a glare free RGB image from a uterine cervix, as shown in figure 1. Glare free imagery can be obtained either by cross-polarized image acquisition or glare removal pre-processing. We used a glare removal algorithm to remove the glare in figure 1b.²

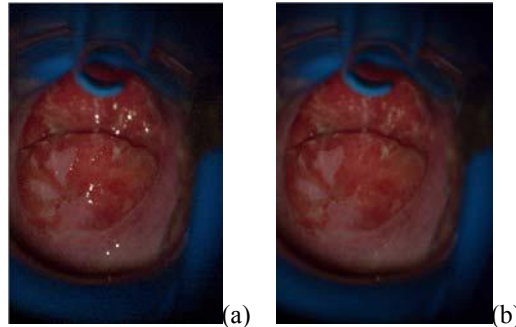


Figure 1: (a) Original uterine cervical image; (b) Glare free RGB image.

We also designed a framework of robust, real-time algorithms to fulfill automatic image quality assessment based on statistical, morphological and signal processing methods (Figure 2). First, the cervical region, used as the processing Region-Of-Interest (ROI), is detected using a hue color cluster that discriminates between cervix and background. Then an adaptive peak-removing histogram equalization algorithm is employed to assess contrast. Next, a novel frequency based method is used to fulfill the out of focus blur evaluation. Finally, the contamination detection algorithm is fulfilled by machine learning and classification algorithms. The architecture of the algorithm framework is flexible and, permits addition and/or substitution of algorithms to extend and improve performance.

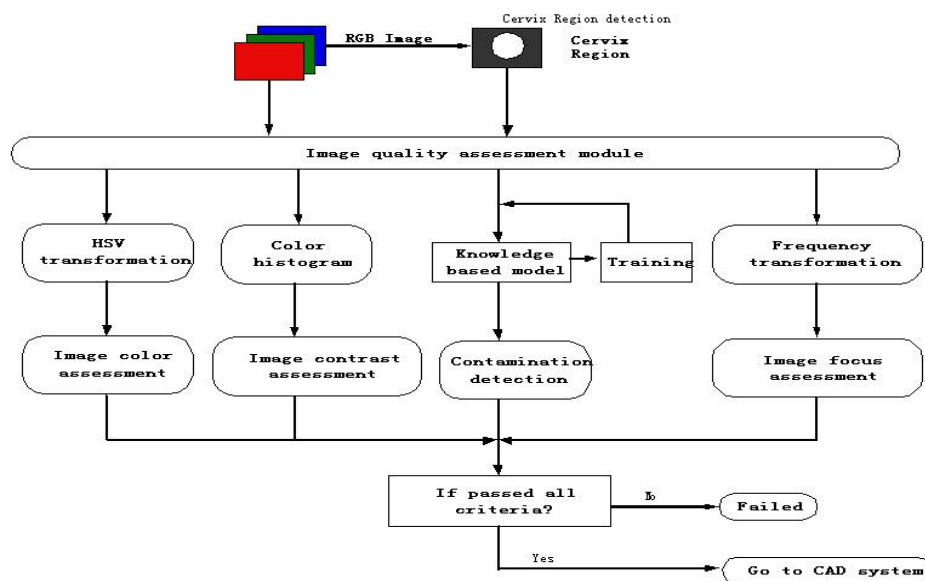


Figure 2: Algorithm framework

2.1 Cervix region detection and region contrast assessment

The cervix region, used as the ROI for further image processing, is detected using a hue color classifier that discriminates between the cervix and the background. First, the Hue-Saturation-Intensity (H-S-I) transformation and the histogram smoothing are performed. Then, an Expectation-Maximization (EM) cluster is employed to fit the two Gaussian peaks in the histogram, and the segmentation is achieved based on the color distribution likelihood. Finally the post-processing is performed. Furthermore, we will show how we performed region and contrast assessment based on cervix region detection results. The algorithm framework of cervix region detection is as follows:

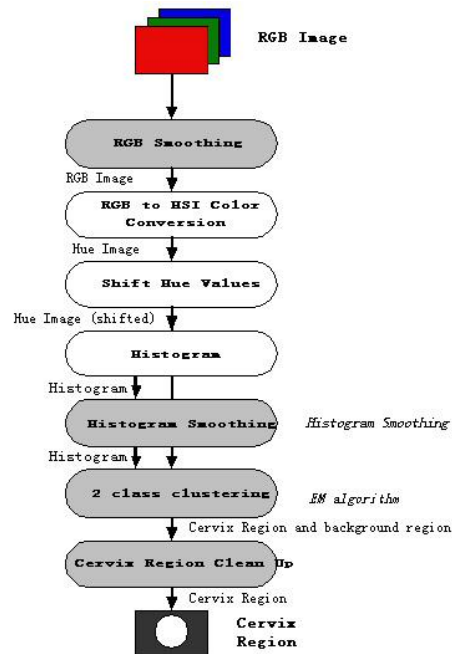


Figure 3: Algorithm framework of probability based cervix region detection

2.1.1 HSI transformation

The hue color feature is used to characterize the color of the pixels. The input RGB images are transformed to the H-S-I color space for the following calculations. The hue values are represented as 8bit values [0 ... 255]. The original RGB image and transformed hue image are shown in figure 4a, figure 4b. The hue values of interest are located at the 255 to 0 roll-over spectrum (red-color) of the 8-bit hue values [0 ... 255], which can easily be seen in the histogram of the hue values in figure 4c.

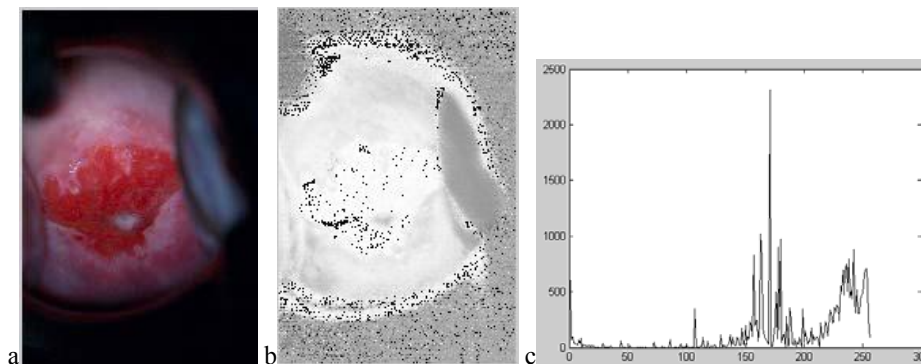


Figure 4: a) Original RGB image, b) transformed hue image and c) histogram of hue values

To simplify the calculations and visualization we shift all hue values by the middle value of the histogram (127). The shifted histogram is shown in figure 5a. As a reference, the red color now corresponds to the hue value of 127. Figure 5b shows the shifted hue image.

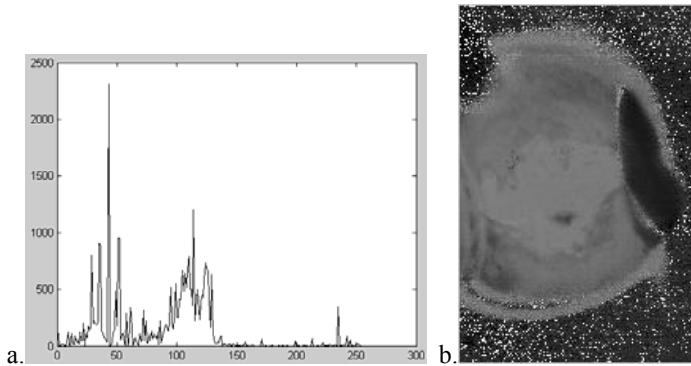


Figure 5: a) Histogram of shifted hue values, b) Shifted hue image.

2.1.2 Histogram smoothing

The hue value histogram is very noisy; therefore, it is important to smooth the histogram in order to get a uniformed distribution. We first smoothed the image in each R, G, B component separately, with a median filter followed by a Gaussian filter. The results on the RGB color image, the hue characteristics image and the histogram are shown in figure 6a, 6b and 6c respectively.

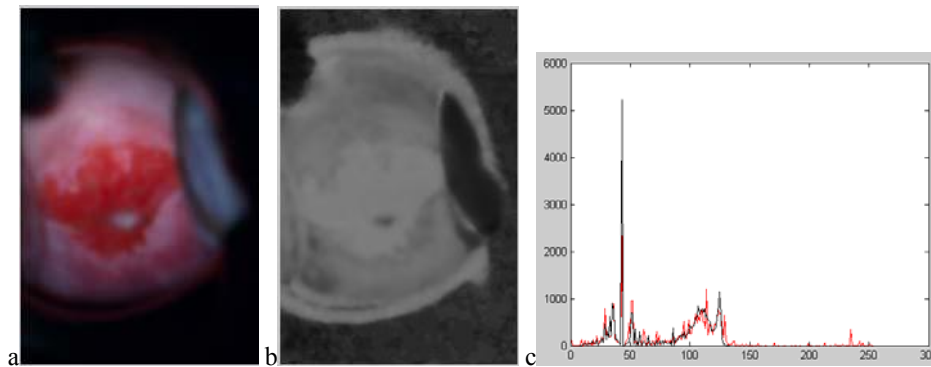


Figure 6: a) RGB color image - RGB smoothed; b) Hue image - RGB smoothed and c) Histogram of hue values - RGB smoothed.

Then the histogram itself is smoothed using an opening-closing Alternating Sequential Filter⁸ (ASF) with a horizontal line structuring element of size 2. Figure 7 shows the resulting smoothed histogram (black) and, for reference the histogram before smoothing (red).

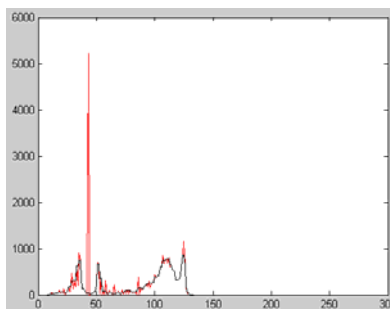


Figure 7: Histogram of hue values – histogram smoothed.

2.1.3 Classification and post-processing

The expected hue value histogram of the scene model has a peak (pink) for the pink region corresponding to the cervix and vagina sidewalls, and a very close peak (brown) to the right of it (hue is an angle) when the body parts outside the cervix and vagina sidewalls are visible, as shown in figure 8. Instead of heuristic based thresholding, we used an EM algorithm⁹ as a probability based method to separate the two peaks by fitting the histogram into two mixture Gaussian models.

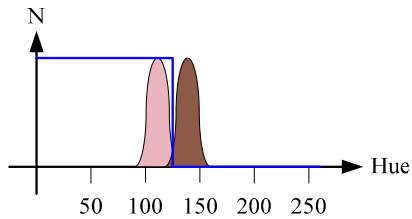


Figure 8: Scene model expected histogram of hue values.

The fitted Gaussian model and the cervix segmentation result are shown in figure 9:

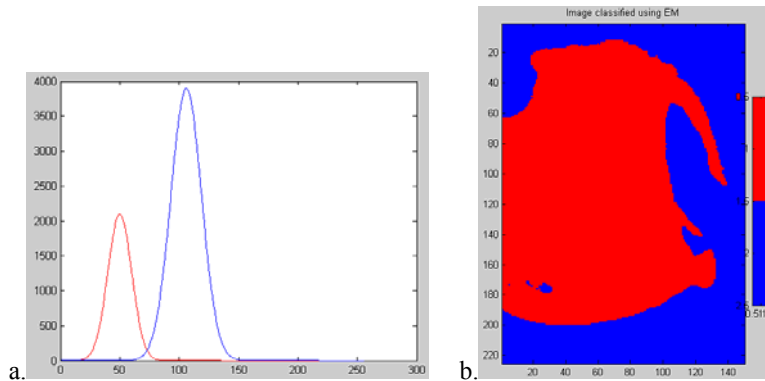


Figure 9: (a) Fitted Gaussian model (2 classes); (b) Segmentation result

After classification, the cervix region needs to be post-processed, as shown in figure 10. First, holes are closed. Then, small regions are deleted using an opening-closing alternative sequential filter with reconstruction using a cross as structuring element.

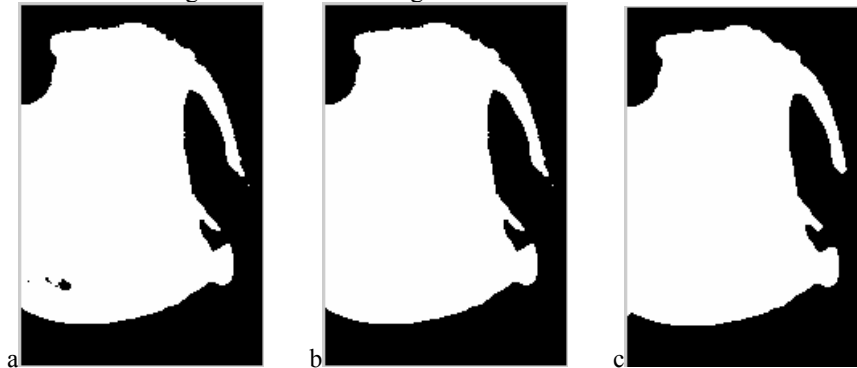


Figure 10: (a) Cervix region – Obtained by classification, (b) Cervix region – closed holes. (c) Cervix region – remove small regions.

2.1.4 Region assessment

After the cervical region of interest has been detected, this information can be used to guide assessment of certain settings that affect image quality. Poor image quality can be attributed to several factors, such as:

1. Incorrect camera zooming: which makes the cervix region too small comparing to the entire image size.
2. Incorrect camera positioning: that makes the cervix region not well located in the center of the image.
3. The presence of foreign objects such as a speculum or cotton swab, which will make some of the cervix region invisible. Also, patient movement may cause portions of the cervical ROI to be excluded.

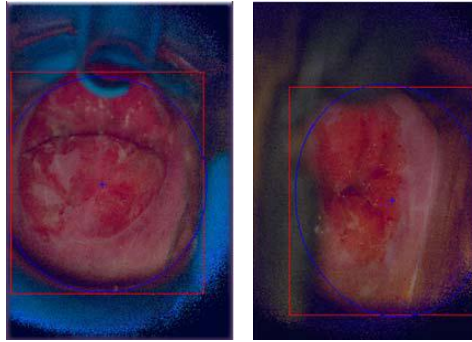
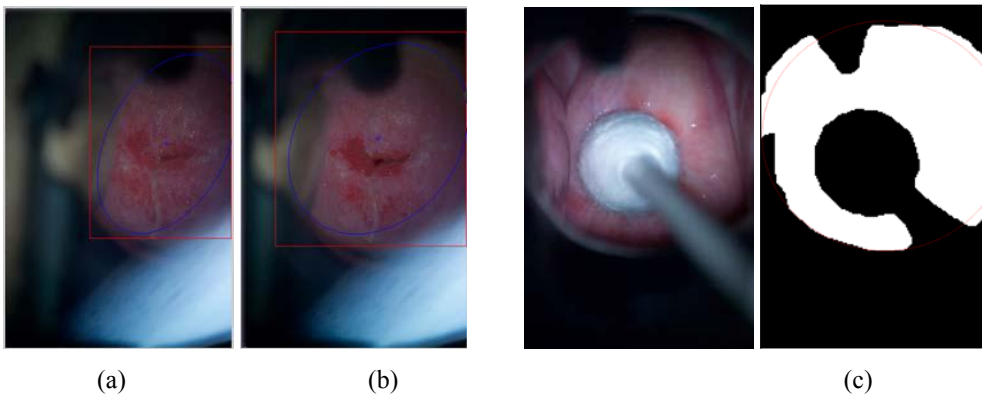


Figure 11: Blue circles represent the ellipse fitting, red rectangles represent the surrounding box and the blue crosses represent the center of mass of the cervix region.

When the above circumstances occur, our system should automatically detect them and block them from entering the subsequent CAD system. Our method consists of the following steps: calculating the surrounding box of the detected cervix region, calculating the mass center and searching for an ellipse fitting of the detected cervix region (see figure 11). All of the above information is used to do region assessment based on the following criteria:

1. For assessment of incorrect camera zooming, we set a threshold as the ratio between cervix region area and entire image area. When the ratio is larger than the threshold, we deem the camera zoom is satisfactory; otherwise an error message will be displayed. (See figure 12a for example)
2. For assessment of incorrect camera positioning, we calculate the distance between the cervix region mass center and the image center. If the distance is smaller than a certain threshold, we deem that the cervix region is well centered. Otherwise the camera positioning is not satisfied and an error message will result. (See figure 12b for example)
3. For assessment of foreign objects or partially visible cervix region, we compare the detected cervix region with the fitted ellipse (based on the assumption that the fully visible cervix region is ellipse or near ellipse). If the area of eroded region is larger than a certain threshold, we deem that the ROI is obscured by foreign objects. (See figure 12c~12e for example)



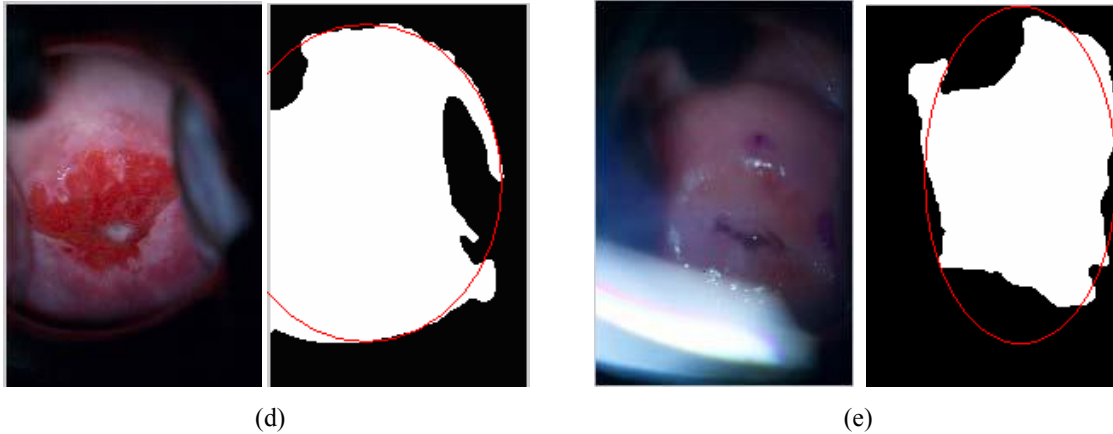


Figure 12: Examples of region assessment results. (a) Incorrect camera zooming, cervical region is too small. (b) Incorrect camera positioning, cervical region is not well centered. (c~e) three examples of foreign objects or partially visible cervical region, c: existence of cotton swab, d) existence of speculum, e) excessive illumination caused by an abnormal camera setting, which impaired the image visibility. All three examples above indicate poor image quality due to incomplete visualization of the cervical region.

2.1.5 Image contrast analysis

The purpose of the image contrast analysis is to assure that the camera settings yield good image contrast. A simple but robust method involves histogram analysis over the region of interest. Empirically, our region of interest is the cervix, which is usually pinkish. We can identify the quality of contrast by analyzing the dynamic range of the red channel. From our experiments, if the peak of the histogram in red channel is in the range of 200 to 250 (note that all the glints have been removed by preprocessing), we may consider the images to have good contrast. Figure 13 shows an image with good contrast. Figure 13 shows an image with good contrast.

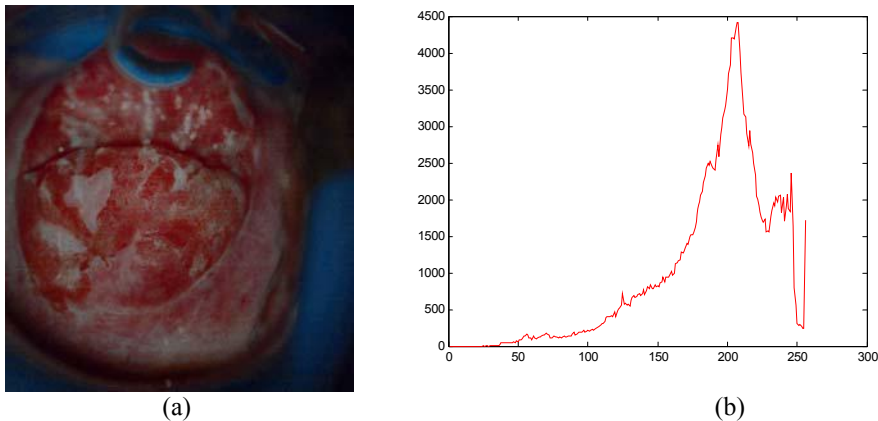


Figure 13: (a) An example of good image contrast, and (b) its histogram in the R Channel.

In real applications, multiple peaks occur frequently. In this situation, we also employ the histogram smoothing and Gaussian fitting, as described in section 2.1.2 and 2.1.3. Thus, we may analyze the dynamic range by the rightmost peak of the R channel histogram, which provides more robust results, as shown in figure 14.

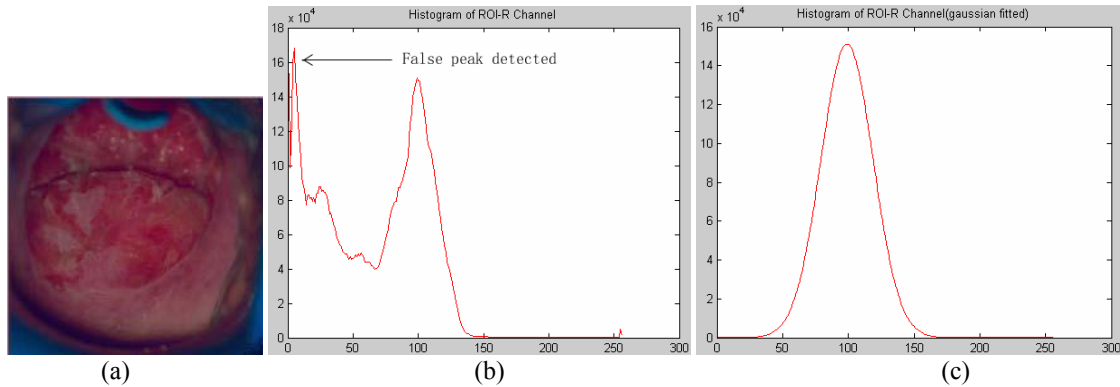


Figure 14: An example of image contrast analysis, (a) Input image; (b) corresponding histogram in R channel (False peak exists); (c) Gaussian fitted Histogram (retain only the rightmost peak)

After the one-peak histogram is obtained by Gaussian fitting, we deem the dynamic range requirement is satisfied if the peak of the histogram is in the range of 200 to 250. Otherwise an error message will be displayed indicating incorrect camera contrast setting.

2.2 Image blur evaluation

Image blur evaluation would be an easy problem to overcome given the availability of a reference image for comparison. In this case, a simple measurement such as entropy or Pixel Signal to Noise Ratio (PSNR) can be used for this purpose. However, in our application, the blur evaluation is in a perceptual base without any reference. Therefore, the selection of appropriate algorithms is limited by the following reasons: a) no reference image. b) various causes of the blur (e.g., camera out of focus, motion blur, or both). A frequency based method makes real-time application possible because of its fast processing speed. However, most frequency-based methods are sensitive to structure changes. An image that contains more structures but looks blurry may actually reflect higher quality than an image that contains fewer structures but has no blur. Thus, we have selected a normalized image power spectrum method as the quality measure. This method has been tested on both simulated images and real cervical images.

The algorithm flow can be described as the following steps, (see figure 15):

1. Divide the image into non-overlapping blocks.
2. For each block, compute local representatives based on frequency information.
3. Compute global statistics from local representatives obtained from Step 2.
4. Determine whether the image is blurred based on the global statistics.

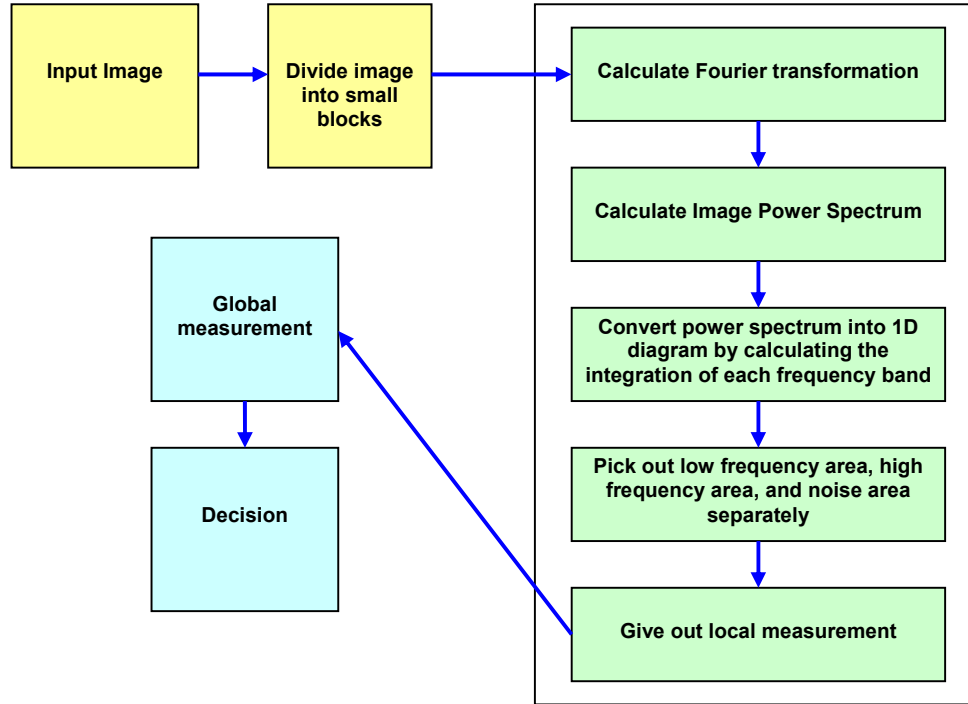


Figure 15: Flowchart of image blur evaluation.

2.2.1. Local representative

The local representative is calculated by the image power spectrum¹⁰, which is generated using the following formula:

$$P(u,v) = \frac{|H(u,v)|^2}{\mu^2} \quad (1)$$

Where $\mu^2 = |H(0,0)|^2$, $H(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} \exp[-2\pi i y \frac{v}{M}] \exp[-2\pi i x \frac{u}{M}] h(x,y)$, $u, v = -\frac{M}{2} \dots \frac{M}{2}$, and $h(x,y)$ is the pixels in the given image block.

After the above formula, the power spectrum has already been normalized by the zero components, which is shown as figure 16a.

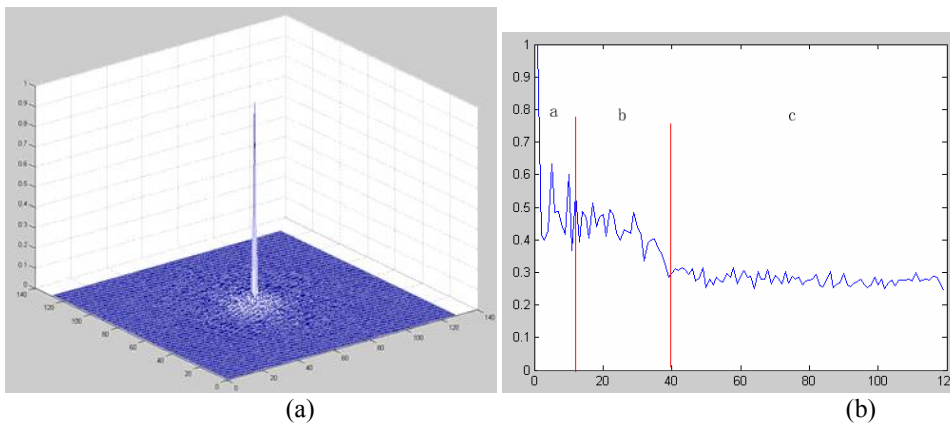


Figure 16: (a) 2D display of image power spectrum; (b) Conversion of image power spectrum into 1D diagram.

Then, this 2D image power spectrum is transformed into a 1D diagram. In order to analyze the energy property in each frequency band, polar coordinate integration is used according to each radial value (See figure 16b). Since there is no reference image to compare with, the only solution is to find some intrinsic information inside the image. Noting that image power spectrum can be a statistical tool for texture analysis and the high frequency information of the texture is always damaged in blurred image, we separate the power spectrum into three parts. We consider that the low frequency area represents structural information invariant to blur, and the high frequency area represents detailed information, which is more sensitive to blur (See figure 16b for example, we separate the image power spectrum 1D diagram into 3 parts: a, b, c, which represent low frequency area, high frequency area and noise area, respectively). By analyzing the ratio between these two integrations, we can calculate the degree of blur (note that the noise spectrum has been discarded by a threshold).

2.2.2. Global statistics

After the blur degree of each small block has been determined, the global measurement is used to analyze the entire image. This can be done by using the percentage of the numbers of blurred blocks occupied in entire image. Furthermore, we added different weights to blocks in the center and blocks in the periphery, since we are more concerned with the image quality in the center. Thus, if the coverage of blurred blocks is less than 25 percent of the entire region of interest, we deem that the image is clear, otherwise an error message will be displayed that indicates that this image is blurry. Figure 17 shows an example of blurred block detection in the region of interest.

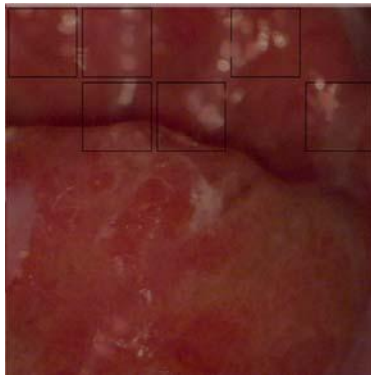


Figure 17: Blur detection result, (black squares represent areas being detected as blur).

2.3 Contamination detection

Contamination detection is a typical classification problem in machine learning. However, in a CAD application, it is heavily dependent on “ground truth” annotations. Ideally, if the annotations are reliable and adequate, all contaminants could be detected after training. In this case, several supervised pattern recognition algorithms can be employed, such as neural network¹¹, support vector machine (SVM)¹² and Bayes net¹³. In this paper, a joint texture/color model is first employed to generate the feature vector for each image pixel¹⁴. Then we use SVM algorithm to detect the contaminants in the cervical images¹⁵.

We have used three color features and three texture features for the training process. The three color components are the L*a*b* coordinates found after spatial averaging using a Gaussian filter, and the three texture components are the anisotropy, polarity and contrast as described by Zhang et al¹². However, in contrast to their approach using an EM algorithm as unsupervised clustering, we utilized a SVM algorithm for supervised classification, based on our adequate, accurate ground truth annotations. We provide here some preliminary results of the blood detection. However, for other kinds of contamination detection such as mucus and purulence, similar algorithms can be designed if the ground truth annotations are accurate. Thus, this is an open problem to which we may add new features in the future. Figure 18 shows some experimental results on blood detection.

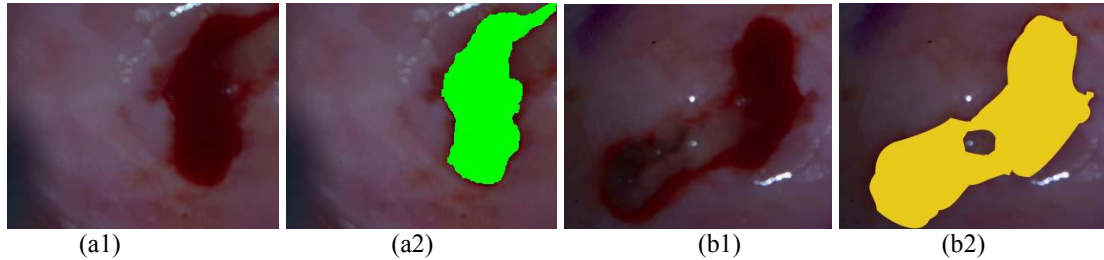


Figure 18: two examples of contamination detection results.
a1: A portion of cervical image, a2: Contamination detection (green area), b1: A portion of cervical image,
b2: Contamination detection (yellow area)

3. DISCUSSION AND CONCLUSIONS

This paper introduces a series of automated algorithms for assessing cervical image quality. We have tested the performance on 149 clinical datasets acquired from multiple clinical sites, with color calibration pre-processing employed¹⁶. Here we give out some of our comments to the performance of these algorithms.

1. The detection of cervical ROI is fully automatic, real-time and parameter free, and gives robust results. Based on the ROI detection, the color analysis, region analysis and contrast analysis can subsequently be performed.
2. The image blur evaluation that we have presented is a novel frequency based method using a single image without any reference, which obtained good results on several datasets. However it is heavily dependent on the selection of thresholds, and the thresholds should be carefully chosen in order to achieve good result. These thresholds vary a lot according to the image content, which is inevitable in case of absence of reference image. Thus there could be two possible improvements: the first is using adjustable focus-length camera in the future. The other solution is to perform this threshold initialization in the camera calibration stage, if possible. Once the proper parameters have been fixed, our algorithm can give out good results.
3. We have shown some preliminary results of contamination detection also. The key idea is using a supervised classification based on accurate ground truth annotations. If the annotation is accurate and represents the range of features commonly encountered, our supervised algorithm can give robust contamination detection results.

As part of a CAD system for colposcopy, several algorithms for automated image quality assessment has been presented that are designed specifically to deal with the digital colposcope application. We have developed a systematic framework for cervical image assessment that includes not only optical analysis but also anatomic analysis.

The future efforts of our image quality assessment include: 1. Design more contamination detection algorithms beside blood detection. 2. Design image quality enhancement algorithms to overcome deficiencies in imagery, such as image contrast enhancement or image deblurring algorithm.

ACKNOWLEDGEMENT

The authors would like to thank Dr. Daron Ferris for his excellent colposcopic image annotations.

REFERENCES

¹ <http://www-depdb.iarc.fr/globocan2002.htm>

- ² Lange H.; Automatic glare removal in reflectance imagery of the uterine cervix; SPIE Medical Imaging 2005; SPIE Proc. 5747, 2005
- ³ Reid R, Scalzi P. Genital warts and cervical cancer. VII An improved colposcopic index for differentiating benign papillomaviral infection from high-grade cervical intraepithelial neoplasia. *Am J Obstet Gynecol* 1985;153:611-618
- ⁴ B.S. Apgar, Brotzman, G.L. and Spitzer, M., *Colposcopy: Principles and Practice*, W.B. Saunders Company: Philadelphia, 2002
- ⁵ Lange H. and Ferris, Daron G.; Computer-Aided-Diagnosis (CAD) for colposcopy; SPIE Medical Imaging 2005; SPIE Proc. 5747, 2005
- ⁶ Gustafsson U., McLaughlin E., Jacobson E., Håkansson J., Troy P., DeWeert M., Pålsson S., Soto Thompson M., Svanberg S., Vaitkuviene A., and Svanberg K.; Fluorescence and reflectance monitoring of human cervical tissue in vivo -a case study; SPIE Photonics West Biomedical Optics (BiOS) 2003; SPIE Proc. 4959, 2003
- ⁷ Gustafsson U., McLaughlin E., Jacobson E., Håkansson J., Troy P., DeWeert M., Pålsson S., Soto Thompson M., Svanberg S., Vaitkuviene A., and Svanberg K.; In vivo fluorescence and reflectance imaging of human cervical tissue; SPIE Medical Imaging 2003; SPIE Proc. 5031, 2003
- ⁸ SDC Morphology Toolbox for MATLAB ® Version 1.3 of 21Apr04 <http://www.mmorph.com/> - SDC Information Systems
- ⁹ Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38.
- ¹⁰ Nill N., Bouzas B., Objective Image quality measure derived from digital image power spectra, *Optical engineering*, April 1992, Vol. 31, 813-825.
- ¹¹ Claude I., Winzenrieth R., Pouletaut P., and Boulanger J., Contour Features for colposcopic image classification by artificial neural networks, in *Proceedings of international conference on Pattern Recognition*, 2002, 771-774.
- ¹² Zhang J., Liu Y., Zhao T., SVM Based Feature Screening Applied to Hierarchical Cervical Cancer Detection, *International Conference on Diagnostic Imaging and Analysis (ICDIA 2002)*, August, 2002.
- ¹³ Friedman, N., Linial, M., Nachman, I., Pe'er, D., Using Bayesian Networks to Analyze Expression Data, *RECOMB 2000*: 127-135 (2000).
- ¹⁴ Carson C., Belongie S., Greenspan H. and Malik J., Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying; *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(8), 1026-1038, August 2002.
- ¹⁵ Chang C. and Lin J. Training nu-support vector regression: theory and algorithms, *Neural Computation*, 14(2002), 1959-1977
- ¹⁶ Li, W., et al, A New Image Calibration Technique for Colposcopic images, *SPIE Medical Imaging 2006*; 2006