Does Multispectral Texture Features Really Improve Cervical Cancer Detection?

Tong Zhao, Jiayong Zhang and Yanxi Liu

Robotics Institute, Carnegie Mellon University
Pittsburgh, PA 15213

ABSTRACT

For cervical cancer detection, the performance of multispectral texture (MST) features extracted from multispectral Pap smear images is evaluated. In this study we carried out pairwise comparisons between different image features, including MST versus average spectral texture features (AST, without spectral information), and MST versus multispectral intensity features (MSI, without texture information). We demonstrate, experimentally, that well-selected MST features combining both multispectral and texture information can achieve better classification results (ROC curves) for cervical cancer detection from multispectral Pap smear images. Furthermore, we investigate which type of wavelet texture features (orthogonal, bi-orthogonal or non-orthogonal) individually or in combination is most effective.

1 Introduction

Two kinds of image features are commonly used on traditional image modalities for cervical cancer detection: namely shape features and texture features [1,2]. Due to the noise, scaling, overlapping of the cells, the shape features extracted directly from the original images are not always robust for automatic classification task. Though recent research on texture analysis has shown that algorithms using the multiresolution wavelet transform features achieve good performance on various type of image classification problems, given a new image modality, there is no intuitive way to know which mother wavelet to choose or to what level the decomposition needs to be taken for a specific problem. We are interested in finding out whether multispectral Pap smear images provide additional information for cervical cancer detection than traditional image modalities. There is little existing work on how to analyze multispectral Pap smear images. Most existing work using multispectral images is carried out for mining, remote sensing and cloud classification, which are targeted at physically quite different materials than what is captured by Pap smear images. In this paper, we study the performances of different wavelet texture features in multispectral Pap smear images and provided a detailed comparison for

1. multispectral texture (MST) features and average spectral texture (AST) features
2. MST features and multispectral intensity (MSI) features;
3. Finally, orthogonal wavelet transform (DB2 and DB16), bi-orthogonal wavelet transform (Bior2.2) and non-orthogonal wavelet transform (Gabor wavelet) features.

2 Texture Feature Extraction

For each non-background pixel in a microscopic Pap smear image, different kinds of image texture features are extracted from a fixed square window centered at the pixel.

2.1 Intensity statistical features

Higher-order ($f(x^n)$, $n \geq 3$) statistical approaches have been successfully used in textural recognition. We apply both low and high order statistical features to extract the cell texture, including maximum, minimum, range, median, mean, standard deviation, energy, skewness, kurtosis and entropy [3].

2.2 Wavelet transform features

The filter coefficients used for computing **orthogonal wavelet transform** are the 2-tap (low order) and 16-tap (high order) Daubechies wavelets (DB2 and DB16) [4], and for the **biorthogonal** case we use 5/3 filters (bior2.2) proposed by Le Gall [5], the bior2.2 is also evaluated by Ma in texture annotation [6]. In our case, the energies of the approximate parts of the cell images are always greater than the other three parts, we adopted therefore has the same structure as the pyramid-structured or tree-structured decomposition[7]. decomposi-tion. We use 2 levels of decomposition of the wavelet transform, which is only performed on the approximation part, therefore 8 subbands are generated in the decompositions. **Gabor wavelets**, as a non-orthogonal filter bank, are expected to extract different texture features than the above wavelets. These texture features are defined as [8]

$$g_{m\theta}(x,y) = \left(\frac{a^{-m}}{2\pi\sigma_1\sigma_2}\right)\exp\left[-\frac{1}{2}\left(\frac{x'^2}{\sigma_1^2}+\frac{y'^2}{\sigma_2^2}\right)+2\pi i W x'\right]$$

where $i=\sqrt{-1}$, σ_1 , σ_2 , W are given, and $x' = a^{-m}(x\cos\theta + y\sin\theta)$, $y' = a^{-m}(-x\sin\theta + y\cos\theta)$. By varying these parameters, a filter can be made to pass any elliptical region of spatial frequencies. Gabor filters are generated according to different *m* ($m=1,...,M$) and θ ($\theta=\theta_1,...,\theta_N$). Considering the performance efficiency and computational complexity, we use 3 scales, 4 orientations to generate 12 Gabor filters.

2.3 Wavelet transform-based features

The wavelet transformation involves filtering and subsampling. A compact representation needs to be derived in the transform domain for classification and detection. The ten statistical features mentioned in Section 2.1 from each of the subbands at each decomposition level are used to construct the feature vector. This results in 10 feature components in each subband, therefore 80 (10x8) features for 2 levels orthogonal and 90 (10x9) biorthogonal wavelet transformations (80 subband features plus 10 statistical features on the original block), and 120 (10x12) features for 12 Gabor filters, totally 370 features for each pixel-window.

2.4 Wavelet-based multispectral texture (MST) features

Using a micro-interferometric spectral imaging setup, a set of *multispetral* Pap smear images is generated containing both normal and cancerous cells. There are total of 52 bands evenly from 400 nm (band 52) to 690 nm (band 1) for each image. Note, this is different from the work by Balas et al, where they use multispectral imaging directly in human body [9]. Since the image spectral quality of the first 20 bands (1 to 20) are much better than the rest of the 52 bands, we extracted the above wavelet texture features on each odd bands of those 20 bands, thus a total of 3700 features is obtained from the 10 spectra.

3 Feature Screening

Not all the image features in this 3700 feature set are useful for cervical cancer detection. We are looking for computationally effective method to rule out irrelevant and redundant feature dimensions in this initial feature set. However, existing feature subset selection algorithms cannot handle a feature set with thousands of dimensions. In order to reduce the feature space dimension we employ sequentially two simple feature quality evaluation measures: **Information Gain** (IG) and **Augmented Variance Ratio** (AVR). IG of a feature X is defined as the difference between the prior uncertainty and the expected posterior uncertainty, while AVR is the ratio of the between-class variance to the within-class variance of the feature, with a penalty for those features with small within-class variance but close inter-class means. Details can be found in [10].

4 Performance Evaluation

4.1 Experimental Setup

The proposed approach for cervical cancer detection in multispectral Pap smear images have been evaluated on a database containing 40 images, with a total of 149 cells (41 cancerous and 108 normal). For each pixel to be classified, various image features are extracted in a 16x16 block centered at the pixel over different bands, thus a very high dimensional MST feature vector is associated with each pixel. We collect a total of 156,732 sample vectors from all 40 images.

Considering the fact that samples from the same image are often highly correlated, we randomly choose the image samples (no pixels from the same image exist in both training and test sets), IG and AVR ranking scores are first computed, and those irrelevant features with low scores are removed from the feature set. Then a modified Gaussian classifier [10] is trained based on the selected features. Finally, the classifier is applied to the samples of the selected features from the test images. By adjusting the decision threshold in the classifier, we record false positive rates (FPR) at different true positive rate (TPR) values. This procedure is repeated 40 times, and the results are averaged.

4.2 Effect of Multispectral Information

To evaluate the performance of multispectra, we compare the MST features against Average Spectral Texture (AST) features. The same intensity and wavelet texture features are generated on the average images of the first 20 bands. Since there is only one band for

the AST features, the feature number of AST is only 1/10 of that of MST. **Table 1** shows the comparative results.

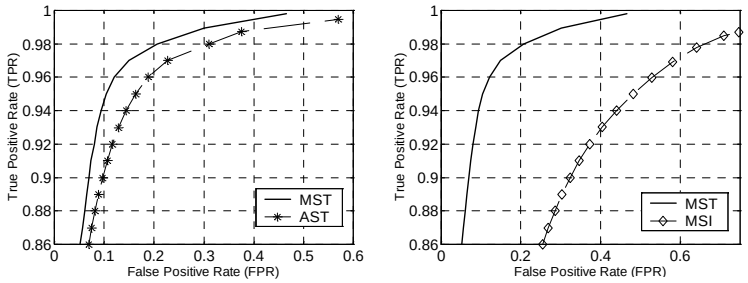
Table 1 Feature selection and classification during the feature screening of MST and AST The values showed in last three rows are FPRs at different TPR values with STD

	# of features	# of screened features	TPR = 0.86	TPR = 0.9	TPR = 0.94
MST	3700	45	0.0442±0.0727	0.0729±0.1096	0.1028±0.1434
AST	370	45	0.0692±0.0505	0.0973±0.0674	0.1440±0.0866

We can see that MST features contain a huge number of irrelevant features, after IG screening, the number of features is decreased to 45. Though the numbers of final selected features are the same for MST and AST, the AVR value of MST is higher than that of AST, this implies the information contained by MST is more discriminative than AST. Figure 1(a) gives the ROC curves of the performance of MST and AST showing MST's superiority.

4.3 Effect of Texture Information

Another aspect worth to compare is the performance of texture features. To evaluate the performance of texture features, the intensity value of each pixel is directly treated as the intensity feature of the corresponding pixel in the odd bands of the top 20 bands. 10 MSI features are obtained from the odd bands of top 20 bands. Figure 1(b) shows that the performances of MSI are much worse than that of MBT. Though we can see from Pap smear images there is a big difference between normal and cancerous cell intensities by human experts, it is hard to classify cells by the single cue of intensity.



(a) **(b)**

Figure 1 The comparative ROC curves (a) MST versus AST (b) MST versus MSI

4.4 Wavelet texture features comparison

Which type of wavelet features is the best to classify multispectral Pap smear images? There is no answer to this important question yet. We choose three basic categories in wavelet transforms: orthogonal, bi-orthogonal and non-orthogonal features for a comparative study in cervical cancer detection.

The wavelet texture features we used are DB2, DB16, Bior2.2 and Gabor filters. Table 2 and Figure 2 shows the feature selection and classification results of the 4 types of wavelet features and their combination features.

Table 2 Performances of combination and the individual kinds of wavelet features

Type of features	DB2	DB16	Bior2.2	Gabor	Combination
# of features x bands	80x10	80x10	90x10	120x10	3700
# of screened features	47	35	41	21	45
TPR = 0.8600	0.0622±0.08	0.0565±0.08	0.0559±0.07	0.0662±0.10	0.0442±0.07
TPR = 0.9000	0.0980±0.12	0.0778±0.11	0.0850±0.10	0.0812±0.12	0.0729±0.11
TPR = 0.9500	0.1667±0.18	0.1180±0.15	0.1459±0.15	0.1098±0.16	0.1138±0.15

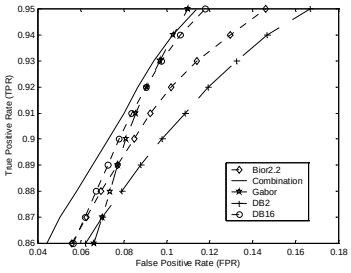


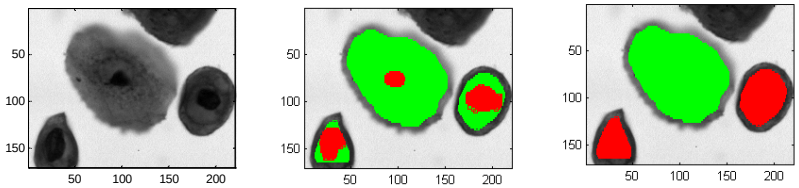
Figure 2 The ROC curves of the combination and 4 kinds of wavelet texture features

To summarize the observations in multispectral Pap smear images for cancer cell detection:

- In general, the combination features can achieve the overall best performance among all individual kinds of wavelet texture features;
- The Gabor and DB16 features have better performance than DB2 and Bior2.2, even though the IG and AVR values of Gabor are far less than others;
- The orthogonal wavelet features are slightly better than the bi-orthogonal ones;
- The high order orthogonal wavelets are better than the low order orthogonal ones.

4.5 Cell level detection

We utilize the shape information in the cell level after the pixel level classification. We use automatic segmentation, morphology smooth, and area ratio judgment to locate the regions with high probability of cancerous cells. Among the 149 cells, only 1 cancer cell was detected as normal cell, and one normal cell has false positive alarms. Figure 3 provides an example, which successfully detects 1 normal cell and two cancerous cells.



(a) **(b)** **(c)**

Figure 3 The classification result, the dark parts are classified as positives and the light ones negatives. (a) The original image with one normal cell in the middle and other two cancerous cells, (b) Pixel level classification, (c) Image level classification

5 Conclusions

We demonstrated, experimentally, that well-selected MST features combining both multispectral and texture information can achieve better classification results (ROC curves) for cervical cancer detection from multispectral Pap smear images. We also compared four types of wavelet transforms for pixel classification, and found that the combined features show the best performance. Although computationally more expensive, Gabor transforms are easy to interpret and offer flexibility in controlling the orientation and scale information, and are amenable for developing scale and orientation invariant features. We are currently investigating new feature selection algorithm, and are trying to combine pixel-level texture features more effectively with high-level shape analysis.

6 Acknowledgement

This research is supported in part by a PA state grant ME #01-738 on ‘‘Cancer Informatics: From Molecules to Clinical Outcomes”, and an award from NIH National Cancer Institute Unconventional Innovation Program N01-CO-07119.

References

[1]. Zhong Li, and Kayvan Najarian, Automated Classification of Pap Smear Tests Using Neural Networks, Proceedings of International Joint Conference on Neural Networks (IJCNN '01), Volume: 4, 2001, pp. 2899 –2901.

[2]. Ross F. Walker, Paul Jackway, Brian Lovell and I. D. Longstaff, Classification of Cervical Cell Nuclei Using Morphological Segmentation and Textural Feature Extraction, Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, 1994, pp. 297 -301

[3]. A. K. Jian. Fundamentals of Digital Image Processing. Prentice Hall, Englewood Cliffs, 1989.

[4]. Ingrid Daubechies, Ten Lectures on Wavelets, SIAM, 1992

[5]. D. Le Gall and A. Tabatabai, Subband coding of digital images using symmetric short kernel filters and arithmetic coding techniques, Proc. ICASSP'88, 1988, pp. 761-765.

[6]. W. Y. Ma and B. S. Manjunath, A comparison study of wavelet transform features for texture image annotation, ICIP'95, pp. 256-259.

[7]. Tianhorng Chang and C.-C. Jay Kuo, Texture Analysis and Classification with Tree-Structured Wavelet Transform, IEEE Trans. on Image Processing, 2(4), 1993, pp. 429-441

[8]. B.S. Manjunath and W.Y. Ma. Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8) , August 1996, pp. 837-842.

[9]. C. Balas, G. Themelis, A. Papadakis, E. Vasgiouraki, A. Argyros, E. Koumantakis, A. Tosca, E. Helidonis, A Novel Hyper-Spectral Imaging System: Application on in-vivo Detection and Grading of Cervical Precancers and of Pigmented Skin Lesions, IEEE Computer Society Workshop on Computer Vision Beyond the Visible Spectrum (CVBVS 2001)

[10]. Y. Liu, T. Zhao and J. Zhang, Learning Multispectral Texture Features for Cervical Cancer Detection, Proceedings of 2002 IEEE International Symposium on Biomedical Imaging: Macro to Nano, July 2002.