TEXTURE-BASED CLASSIFICATION OF HYSTEROSCOPY IMAGES OF THE ENDOMETRIUM

M. S. Neofytou[1,4], *Student Member, IEEE,* M. S. Pattichis[3], *Member, IEEE,* C. S. Pattichis[1], *Member, IEEE,* V. Tanos[2], E. C. Kyriacou[1], *Member, IEEE,* D. D. Koutsouris[4], *Member, IEEE,*

Abstract—**The objective of this study was to classify hysteroscopy images of the endometrium based on texture analysis for the early detection of gynaecological cancer. A total of 418 Regions of Interest (ROIs) were extracted (209 normal and 209 abnormal) from 40 subjects. Images were gamma corrected and were converted to gray scale. The following texture features were extracted: (i) Statistical Features, (ii) Spatial Gray Level Dependence Matrices (SGLDM), and (iii) Gray level difference statistics (GLDS). The PNN and SVM neural network classifiers were also investigated for classifying normal and abnormal ROIs. Results show that there is significant difference (using Wilcoxon Rank Sum Test at a=0.05) between the texture features of normal and abnormal ROIs for both the gamma corrected and uncorrected images. Abnormal ROIs had lower gray scale median and homogeneity values, and higher entropy and contrast values when compared to the normal ROIs. The highest percentage of correct classifications score was 77% and was achieved for the SVM models trained with the SF and GLDS features. Concluding, texture features provide useful information differentiating between normal and abnormal ROIs of the endometrium.**

I. INTRODUCTION

Cancer diseases are the second cause of death in USA with over 557.271 deaths in 2002. Gynaecological cancer is the second cause of death among the female population and in 2005 over 40.880 new cases is expected to be diagnosed with an estimation of 7.310 deaths from gynaecological caner [1]. In this study, hysteroscopy imaging is investigated for the assessment of endometrium tissue. Hysteroscopy is considered to be the golden standard technique for the diagnosis of intrauterine pathology [2]. The physician guides the telescope connected to a camera inside the uterus in order to investigate suspicious lesions of cancer [3]. Sometimes, areas of endometrial cancer can be missed or unnecessary endometrial biopsies can be performed. The objective of this study was to evaluate the usefulness of texture analysis for differentiating between normal and

abnormal ROIs from hysteroscopy images of the endometrium for the early detection of gynaecological cancer. To the best of our knowledge, no similar study was carried out for hysteroscopy imaging of the endometrium. In previous work, we used the abdominal cavity of a chicken and calf uterus as experimental tissue, comparing texture feature variability under different viewing conditions such as different angles (with 5 degrees difference) and different distances (of 3 cm in close up view and 5 cm in panoramic view) from the ROI under investigation [4]. The results indicated that for small consecutive angles there is no significant difference in texture features analysis but there is significant difference when comparing panoramic vs. close up views. In another study, we demonstrated that some texture features can be used to differentiate between normal and abnormal endometrium images captured during hysteroscopy [5]. However, in both [4] and [5], the images were not gamma corrected. Instead, image intensity normalization was carried out by white balancing, adjusting the camera gain. This approach introduces same sources of variability in the acquisition. Also, in previous work [6] we proposed a standardized protocol for capturing hysteroscopy images using gamma correction as a preprocessing step.

The significance of color calibration of the CCD camera was also proposed by [7] and [8]. Scarcanski et al. [8] proposed a hardware calibration protocol of hysteroscopy images with very good results according to the experts.

Moreover, the use of color information for the content based retrieval of endoscopy images was also applied successfully by Shunren et al. [9].

In this paper, in sections II, III and IV we present the methodology, results and concluding remarks respectively.

II. METHODOLOGY

A. Recording of Video

The CIRCON IP4.1 [10] medical camera was used. The analog output signal of the camera (PAL 475 horizontal lines) was digitized at 720x576 pixels using 24 bits color at 25 frames per second, and was then saved in AVI format. The Digital Video Creator 120 frame grabber was used [11].

B. Material

A total of 418 hysteroscopy images from the endometrium were recorded from 40 subjects (see Fig. 1). Regions of Interest of 64x64 pixels were cropped and classified into two categories: (i) normal (N=209) ROIs and (ii) abnormal (N=209) ROIs based on the physician opinion and the histopathological examination (see Fig. 2).

C. Gamma correction

Most of the cameras have a nonlinear relationship between the signal voltage and the light intensity. The light intensity input to the medical camera or the output of the display is proportional to the voltage raised to the power gamma.

Gamma is computed using calibration color images, and then applying a non linear fitting. Gamma is computed for the red, green and blue channels, and typical values used by our group based on previous findings were in the range of 1.4 to 1.6 [6].
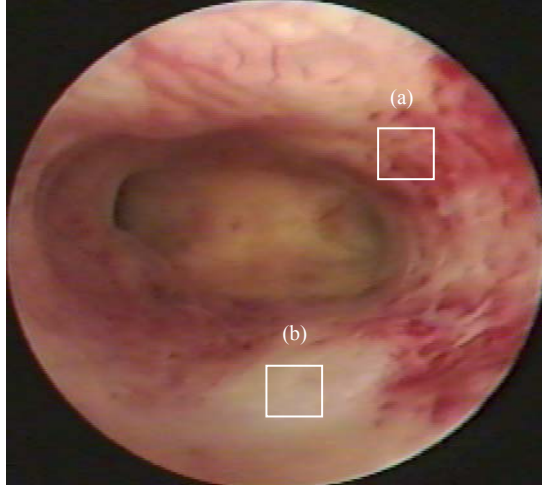

Fig. 1. A hysteroscopy image of the endometrium before gamma correction with ROIs selected by the gynaecologist: (a) normal ROI, and (b) abnormal ROI.

D. Feature Extraction

ROIs were transformed into grayscale using the equation (Intensity=0.299*red+0.587*green+0.114*blue channels) and the following texture features [12] were computed:

Statistical Features (SF): SF features describe the gray level histogram distribution without considering spatial independence. The following texture features were computed: 1) Mean, 2) Variance and 3) Entropy.

Spatial Gray Level Dependence Matrices (SGLDM): The spatial gray level dependence matrices as proposed by Haralick et al. [13] are based on the estimation of the second-order joint conditional probability density functions that two pixels (k, l) and (m, n) with distance d in direction specified by the angle θ, have intensities of gray level i and gray level j. Based on the estimated probability density functions, the following four texture measures out of the 13 proposed by Haralick et al. [13] were computed: 1) Contrast, 2) Correlation, 3) Homogeneity, 4) Entropy. For a chosen distance d (in this work $d=1$ was used), and for angles $\theta = 0^o, 45^o, 90^o$ and 135° we computed four values for each of the above texture measures. The above features were calculated for displacements δ=(0,1), (1,1), (1,0), (1,-1), where $\delta \equiv (\Delta x, \Delta y)$, and their range of values were computed.

Gray level difference statistics (GLDS): The GLDS algorithm [14], [15] is based on the assumption that useful texture information can be extracted using first order statistics of an image. The algorithm is based on the estimation of the probability density p_δ of image pixel pairs

at a given distance $\delta=(\Delta_\chi, \Delta_\upsilon)$, having a certain absolute gray level difference value. Let p_δ be the probability density of $f_\delta(x,y)$. If there are m gray levels, this has the form of an m-dimensional vector whose ith component is the probability that $f_\delta(x,y)$ will have value i. If the picture f is discrete, it is easy to compute p_δ by counting the number of times each
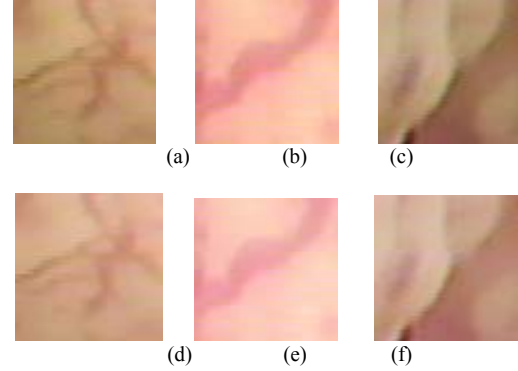

Fig. 2. ROIs of the endometrium. Original images: (a) normal tissue (b) and (c) abnormal tissue. Corresponding gamma corrected images (d) normal tissue (e) and (f) abnormal tissue

value of $f_\delta(x,y)$ occurs, where Δ_χ and Δ_y are integers. Coarse texture images, result in low gray level difference values, whereas, fine texture images result interpixel gray level differences with great variances. Variable i is two pixels gray level difference, m is the number of gray levels and p_δ are the individual probabilities. Features were estimated for the following distances δ=(d,0), (d,d), (-d,d), (0,d). A good way to analyze texture coarseness is to compute, for various magnitudes of δ, some measure of the spread of values in p_δ away from the origin.

E. Image Classification

The diagnostic performance of the texture features was evaluated with two different classifiers: the Probabilistic Neural Network (PNN), and the Support Vector Machine (SVM). These classifiers were trained to classify the texture features into two classes: i) normal ROIs or ii) abnormal ROIs. The PNN [16] classifier basically is a kind of Radial Basis Function (RBF) network suitable for classification problems. This classifier was investigated for several spread radius in order to identify the best for the current problem. The SVM network was investigated using Gaussian Radial Basis Function (RBF) kernels; this was decided as the rest of the kernel functions could not achieve so good results. The SVM with RBF kernel was investigated using 10-fold cross validation in order to identify the best parameters such as spread of the RBF kernels [17]. The leave-one-out method was used for validating all the classification models. A total of 418 runs were carried out for training the classifiers, and the performance of the classifiers was evaluated on the remaining one subset. The performance of the classifier systems were measured using the parameters of the receiver operating characteristic (ROC) curves [17]: true positives (TP), false positives (FP), false negatives (FN), true negatives (TN), sensitivity (SE), specificity (SP), and

precision (PR). We also computed the percentage of correct classifications ratio (%CC) based on the correctly and incorrectly classified cases. The time performance of the PNN was 60 sec and the time performance for the SVM algorithm was 10 min.

III. RESULTS

Table I presents the texture features results before and after gamma correction. The non-parametric Wilcoxon rank sum test was used to decide if there is a significant difference between normal and abnormal ROIs at a=5%. The results indicate that there is a significant difference in texture feature values between normal and abnormal ROIs of the endometrium, both between the original images, and the gamma corrected images. Also, we note that gamma-corrected images were judged to be visually better, as evaluated by the physician. Most importantly, as we can see in Table I, there is a reduction in variance after gamma correction while the entropy values are preserved. This result indicates that gamma correction can be used to reduce the variance without sacrificing the effective range of pixel values (as measured by the entropy). In Table I, we also report the *median*, and the spread of the data (*inter quartile range* (*IQR*) (difference between the 25th and 75th percentile) for all the texture features investigated. It is clearly shown in Table I that the median values for the features SF: *median*, and SGLDM: *contrast*, decrease and increase respectively between normal and abnormal ROIs. Also, Table I presents the statistical analysis between normal ROIs before and after gamma correction, and the statistical analysis between abnormal ROIs before and after correction. We can see that, in all features except one, there is significant difference, leading us to the conclusion that the gamma correction affects significantly the ROI images.

Table II presents the performance of the different PNN and SVM classification models investigated using the texture features. It is clearly shown that the SVM classifier performed better than the PNN classifier. For the SVM classifier the best performance was achieved with the SF+SGLDS followed by the SGLDM+GLDS, and SGLDS with a percentage of correct classifications (%CC) of 77%, 76%, and 76%, respectively. Similar classification performance to the corrected ROIs was also obtained for all models for the uncorrected ROIs. Moreover, similar performance for all models as given in Table II was obtained when the feature sets were transformed using PCA.

IV. CONCLUDING REMARKS

In this study we classified ROIs of the endometrium from hysteroscopy images based on texture features before and after gamma correction. Results showed that the gamma corrected color images were visually better than the originals according to the gynaecologist opinion. There was a significant difference in the SF, SGLDM, and GLDS features investigated between the normal and abnormal ROIs. The highest percentage of correct classifications score was 77% and was achieved for the SVM classifier for the SF+GLDS feature sets. These results support the application of texture analysis for the assessment of difficult cases of normal and abnormal ROIs for gynaecological cancer. the gynaecologist during the operation so as to identify suspicious ROIs for further histopathological examination.

Future work will also investigate multiscale color texture analysis in differentiating between normal and abnormal ROIs of the endometrium.

REFERENCES

[1] Web link: http://www.cancer.org
[2] J.A. Fayez, M.F. Vogel, "Comparision of different treatment methods of endometriomas by laparoscopy," *Obstet. Gynecol.*, Vol. 78, pp. 660-665, 1991.
[3] R. Wenzl, R. Lehner, U. Vry, N. Pateisky, P. Sevelda, P. Husslein, "Three-dimensional video-hysteroscopy: clinical use in gynaecological laparoscopy," *Lancet*, Vol. 344, pp. 1621–1622, 1994.
[4] M.S. Neophytou, C.S. Pattichis, M.S. Pattichis, V. Tanos, E. Kyriacou, D. Koutsouris, "Multiscale Texture Feature Variability Analysis in Hysteroscopy Imaging Under Different Viewing Positions," *Proc. II EFOMP Med. Conf. on Med. Phys.*, Limassol, Cyprus, 2004.
[5] M.S. Neophytou, C.S. Pattichis, M.S. Pattichis, V. Tanos, E. Kyriacou, S. Pavlopoulos, D. Koutsouris, "Texture Analysis of the Endometrium During Hysteroscopy: Preliminary Results," *Proc. IEEE EMBS 04 Conf.*, San Francisco, USA, Sep. 2004.
[6] M.S. Neophytou, C.S. Pattichis, M.S. Pattichis, V. Tanos, E. Kyriacou, D. Koutsouris, "The Effect of Color Correction of Hysteroscopy Images for Quantitative Analysis in Endometrium,"27th Annual International conference of the IEEE engineering in Medicine and Biology Society, 1-4 September, Shanghai, China, 4 pages, 2005.
[7] S. Sheraizin, V. Sheraizin, "Endoscopy Imaging Intelligent Contrast Improvement" 27th Annual International conference of the IEEE engineering in Medicine and Biology Society, 1-4 September, Shanghai, China, 4 pages, 2005.
[8] J. Scarcanski, W. Gaviao, S. Cunha, F. Joao, "Diagnostic Hysteroscopy Video Summarization and Browsing", 27th Annual International conference of the IEEE engineering in Medicine and Biology Society, 1-4 September, Shanghai, China, 4 pages, 2005
[9] X. Shunren, M. Weirong, W. Xiaoying, Z. Zanchao, "A content-Based retrieval system for Endoscopic Images", 27th Annual International conference of the IEEE engineering in Medicine and Biology Society, 1-4 September, Shanghai, China, 4 pages, 2005.
[10] Web link: http://www.acmicorp.com
[11] Web link: http://www.pinnaclesys.com
[12] P.M. Tjoa, M.S. Krishnan, "Feature extraction for the analysis of colon status from the endoscopic images," *BioMedical Engineering OnLine*, Apr. 2003. Available: http://www.biomedical-engineering-online.com/content/2/1/9.
[13] R.M. Haralick, K. Shanmugam, I. Dinstein, "Texture Features for Image Classification," *IEEE Trans. on Systems, Man., and Cybernetics*, Vol. SMC-3, pp. 610-621, Nov. 1973.
[14] C.M. Wu, Y.C. Chen, and K.S. Hsieh, "Texture features for classification of ultrasonic liver images," *IEEE Trans. Med. Imaging*, vol. 11, pp.141-152, 1992.
[15] J.S. Wenska, C.R. Dryer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cyber.*, vol. SMC-6, pp. 269-285, 1976.
[16] D.F Specht, Probabilistic Neural Networks, *INNS Neural Networks* 3(1), 109-118, 1990.
[17] T. Joachims, "Making large-scale svm learning practical. advances in kernel methods - support vector learning". In B. Schölkopf, C. Burges, and A. Smola, editors, *Advances in Kernel Methods: Support Vector Machines*, pages 169-184. MIT Press, Cambridge, MA, USA, 1999.
[18] J. A. Swets and R. M. Pickett, "Evaluation of Diagnostic Systems", Methods from Signal Detection Theory. New York: Academic, 1992.

TABLE I
TEXTURE FEATURES AND STATISTICAL ANALYSIS OF NORMAL (N=209) vs ABNORMAL (N=209) ROIs OF THE ENDOMETRIUM EXTRACTED FROM 40 SUBJECTS. STATISTICAL ANALYSIS WAS CARRIED OUT BEFORE AND AFTER GAMMA CORRECTION BUT ALSO BETWEEN THE NORMAL/ABNORMAL ROIs BEFORE AND AFTER GAMMA CORRECTION. AT a=5%.

| | Original Images | | | | | Corrected Images | | | | | Orig. Vs Cor. Im. | |
| | Normal Cases | | Abnormal Cases | | | Normal Cases | | Abnormal Cases | | | Nor. | Abnor. |
	Median	IQR	Median	IQR	H	Median	IQR	Median	IQR	H	H	H
SF												
Mean	136.802	20.173	124.995	22.056	1	156.061	17.734	144.645	20.553	1	1	1
Variance	67.252	59.666	142.849	102.465	1	54.625	49.251	124.386	78.216	1	1	1
Median	136.96	20.384	123.211	23.34	1	156.435	17.796	143.749	21.756	1	1	1
Energy	0.038	0.013	0.025	0.007	1	0.041	0.014	0.028	0.008	1	1	1
Entropy	3.444	0.334	3.814	0.27	1	3.339	0.332	3.739	0.279	1	1	1
SGLDM												
ASM	0.006	0.003	0.004	0.001	1	0.007	0.003	0.004	0.002	1	1	1
Contrast	4.576	1.152	8.163	4.252	1	3.815	0.886	7.042	3.086	1	1	1
Correlation	0.963	0.022	0.97	0.016	1	0.962	0.024	0.97	0.016	1	1	0
Variance	65.931	59.195	140.559	99.319	1	53.965	48.791	120.847	77.925	1	1	1
Homogeneity	0.454	0.03	0.394	0.041	1	0.479	0.025	0.415	0.039	1	1	1
Entropy	5.48	0.449	6.036	0.398	1	5.309	0.426	5.93	0.395	1	1	1
GLDS												
Homogeneity	0.454	0.029	0.395	0.041	1	0.48	0.025	0.416	0.039	1	1	1
Contrast	4.567	1.149	8.145	4.242	1	3.808	0.884	7.028	3.079	1	1	1
Energy	0.24	0.021	0.197	0.03	1	0.255	0.019	0.209	0.03	1	1	1
Entropy	1.609	0.105	1.841	0.192	1	1.54	0.093	1.775	0.165	1	1	1
Mean	1.581	0.177	2.034	0.427	1	1.445	0.151	1.89	0.349	1	1	1

TABLE II
CLASSIFICATION PERFORMANCE OF THE SVM AND PNN MODELS BASED ON TEXTURE FEATURES FROM THE GAMMA CORRECTED (UNCORRECTED) ROIs. MODELS WERE DEVELOPED TO CLASSIFY NORMAL AND ABNORMAL ROIs OF THE ENDOMETRIUM.

| | Corrected Images (Uncorrected Images) | | | | | | | | | |
	Normal[1]		Abnormal[2]		%CC	%FP	%FN	%SE	%SP	%PR
SVM classifier										
SF	151 (152)	58 (57)	144 (142)	65 (67)	71 (70)	28 (27)	31 (32)	69 (67)	72 (72)	71 (71)
SGLDM	161 (164)	48 (45)	157 (151)	52 (58)	76 (75)	23 (21)	24 (27)	75 (72)	77 (78)	77 (77)
GLDS	166 (166)	43 (43)	137 (137)	72 (72)	72 (72)	21 (20)	34 (34)	66 (65)	79 (79)	76 (76)
SF+SGLDM+GLDS	173 (162)	36 (47)	138 (160)	71 (49)	74 (**77**)	17 (22)	34 (23)	66 (76)	83 (77)	79 (77)
SGLDM+SF	163 (163)	46 (46)	146 (146)	63 (63)	74 (73)	22 (22)	30 (30)	70 (69)	78 (77)	76 (76)
SGLDM+GLDS	162 (161)	47 (48)	155 (158)	54 (51)	76 (76)	22 (22)	26 (24)	74 (75)	78 (77)	77 (76)
SF+GLDS	160 (166)	49 (43)	160 (155)	49 (54)	**77** (76)	23 (20)	23 (25)	77 (74)	77 (79)	77 (78)
PNN classifier										
SF	152 (152)	57 (57)	116 (116)	93 (93)	64 (64)	27 (27)	44 (44)	56 (56)	73 (73)	67 (67)
SGLDM	151 (151)	58 (58)	131 (131)	78 (72)	67 (67)	28 (28)	38 (38)	63 (63)	72 (72)	69 (69)
GLDS	172 (172)	37 (37)	107 (107)	102 (102)	67 (67)	18 (18)	49 (49)	51 (51)	82 (82)	74 (74)
SF+SGLDM+GLDS	165 (165)	44 (44)	136 (136)	73 (73)	72 (72)	21 (21)	35 (35)	65 (65)	79 (79)	76 (76)
SGLDM+SF	156 (156)	53 (53)	131 (131)	78 (78)	69 (69)	25 (25)	37 (37)	63 (64)	75 (75)	71 (71)
SGLDM+GLDS	164 (168)	45 (45)	132 (132)	77 (77)	71 (71)	22 (22)	37 (37)	63 (63)	78 (78)	75 (75)
SF+GLDS	169 (168)	40 (41)	125 (125)	84 (84)	70 (70)	19 (19)	40 (40)	60 (60)	81 (81)	76 (76)

[1]The first column presents the numbers of corrected normal (uncorrected normal) cases that were recognized by the classifier and the second column presents the corrected normal (uncorrected normal) cases that were not recognized by the classifier.
[2]The third column presents the numbers of corrected abnormal (uncorrected abnormal) cases that were recognized by the classifier and the fourth column presents the corrected abnormal (uncorrected abnormal) cases that were not recognized by the classifier.