Texture analysis of tissues in Gleason grading of prostate cancer

Eleni Alexandratou*[a], Dido Yova[a], Dimitris Gorpas[a], Petros Maragos[b], George Agrogiannis[c], Nikolaos Kavantzas[c]

a National Technical University of Athens, Department of Electrical Engineering & Computing, Applied Biophysics and Biomedical Optics Laboratory, 157 73 Zografou Campus, Athens, Greece;
b National Technical University of Athens, Department of Electrical Engineering & Computing, Computer Vision and Speech Laboratory, 157 73 Zografou Campus, Athens, Greece,
c University of Athens, Medical School, Department of Pathology, 75 M.Asias str., 11527, Athens, Greece

ABSTRACT

Prostate cancer is a common malignancy among maturing men and the second leading cause of cancer death in USA. Histopathological grading of prostate cancer is based on tissue structural abnormalities. Gleason grading system is the gold standard and is based on the organization features of prostatic glands. Although Gleason score has contributed on cancer prognosis and on treatment planning, its accuracy is about 58%, with this percentage to be lower in GG2, GG3 and GG5 grading. On the other hand it is strongly affected by "inter- and intra observer variations", making the whole process very subjective. Therefore, there is need for the development of grading tools based on imaging and computer vision techniques for a more accurate prostate cancer prognosis.

The aim of this paper is the development of a novel method for objective grading of biopsy specimen in order to support histopathological prognosis of the tumor. This new method is based on texture analysis techniques, and particularly on Gray Level Co-occurrence Matrix (GLCM) that estimates image properties related to second order statistics. Histopathological images of prostate cancer, from Gleason grade2 to Gleason grade 5, were acquired and subjected to image texture analysis. Thirteen texture characteristics were calculated from this matrix as they were proposed by Haralick. Using stepwise variable selection, a subset of four characteristics were selected and used for the description and classification of each image field. The selected characteristics profile was used for grading the specimen with the multiparameter statistical method of multiple logistic discrimination analysis. The subset of these characteristics provided 87% correct grading of the specimens. The addition of any of the remaining characteristics did not improve significantly the diagnostic ability of the method. This study demonstrated that texture analysis techniques could provide valuable grading decision support to the pathologists, concerning prostate cancer prognosis.

Keywords: Texture analysis techniques, Gray Level Co-occurrence Matrix (GLCM), prostate cancer, Gleason scoring

1. INTRODUCTION

According to the National Cancer Institute, prostate cancer is the second most common type of cancer and the second leading cause of cancer death among men in the U.S.A.[1]. The standard method to assess histological grading of prostate cancer is Gleason score. The Gleason grading system is based on the degree of architectural differentiation. A primary pattern is assigned for the dominant grade and a secondary pattern for the non-dominant grade; the Gleason score is obtained by adding these two values[2]. Recently the College of American Pathologists Conference on solid tumor prognostic factors in conjunction with the World Health Organization Second International Consultation on Prostate Cancer, issued a formal statement regarding a significant modification to the Gleason score (GS)[3]. In brief, they suggested that the Gleason score should be composed of the sum of the grades of the most prevalent component and the most aggressive component.

Although Gleason grading is not the only histological grading system for prostate cancer, it is widely accepted because it takes into account the inherent heterogeneity of prostate cancer, and it has clearly shown that this method is of great prognostic value. The clinical outcome of prostate cancer is closely related to tumor grade and stage. Accurate Gleason

Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues VI,
edited by Daniel L. Farkas, Dan V. Nicolau, Robert C. Leif,
Proc. of SPIE Vol. 6859, 685904, (2008) · 1605-7422/08/$18 · doi: 10.1117/12.763377

score is an independent predictor of tumor prognosis, and hence it has great influence on treatment planning. It requires training in the recognition of the basic patterns and constant application to a critical number of new cases to maintain an acceptable level of accuracy in accordance with international consensus[4].

Gleason grading is based on recognition of organization features of prostatic glands. It is a five level scoring scheme describing the tissue *differentiation*, which is the ability of the tumor to mimic normal gland architecture. According to Gleason grading a tissue can be characterized from grade 1 which is very well differentiated one to grade 5 which is very poorly differentiated and aggressively malignant. Tumors of grades 1 and 2 closely resemble normal prostate and are composed by very pale glands which grow closely together. Grade 3 tumors have also glands, which vary in size and shape more than those in grades 1 and 2 and the invasion of glands into the stroma (muscle) is very prominent. Tumors of grade 4 are identified almost entirely by loss of the ability to form individual, separate gland units, each with its separate lumen. Grade 5 is assigned to tumors when there is no evidence of any attempt to form gland units and only separate cells are visible.

Although Gleason grading is a common method used by pathologists to determine the aggressivity of prostate cancer on the basis of biopsy specimens, the overall accuracy of GS to predict pathologic findings accounts to 58 %, as it is highly dependent on the skill and experience of the pathologist and is a subject to high variability in both intrapathologist and interpathologist reading[5, 6]. There is also an uncertainty to assign it to intermediate disease stage groups and it is not possible to assign to a minute focus of carcinoma on prostate biopsy[7,8].

According to NCI Prostate Portfolio Review Group and the NCI Imaging Sciences Working Group there is urgent need to be developed improved imaging methods for staging and determination of the disease aggressiveness[9]. Taking into consideration that biopsy is the gold standard for cancer diagnosis and staging and that the number of biopsies should be kept to minimum, the introduction of novel imaging and machine vision techniques for acquisition, processing and analysis of images from biopsies is of special importance[10, 11]. However, very little work has been done mainly due to the complexity of histopathological images. Therefore there is the need to develop a computer aided method for more accurate and objective Gleason grading of histopathological tissues.

Diamond et al. have used morphometric and texture features of prostatic tissues in order to distinguish among stromal, normal, and PCa regions of the prostate gland with a diagnostic accuracy of 79%[12]. Roula et al.[13] applied multispectral imaging and achieved an accuracy of 94% for classification of tissues in stroma, benign prostatic hyperplasia, prostatic intraepithelial neoplasia and prostate carcinoma. Texture characteristics of tissue images have been used by Smith et al. for image classification to Gleason grades 1 through 3 and the combined 4 and 5, reporting an accuracy of 90%[14]. Stotzka et al. used structural and statistical features of the spatial distribution of epithelial nuclei in order to distinguish between moderetaly and poorly differentiated samples[15]. In another work of Tabesh et al. color, texture and morphometric cues were used to classify prostate tissues into low and high grade classes.All the above mentioned references are used either for prostate cancer diagnosis or for tissue discriminator in low or high grade. There is no reference concerning prostate cancer classification to the five different grades of the Gleason grading scheme.

In this paper, a new method for objective grading of biopsy specimen was developed in order to support histopathological prognosis. This automated classification method of prostate cancer biopsies was based on texture analysis techniques and more specific on Haralick texture characteristics of microscopic images measured from the Gray level Co-occurrence Matrix (GLCM). Special algorithms for texture analysis and calculation of Haralick texture characteristics were developed in Matlab 7.1. More specific thirteen texture characteristics were calculated from GLCM as they were proposed by Haralick. Using stepwise variable selection, a subset of four characteristics were selected and used for the description and classification of each image field. The selected grading profile was used for image and subsequently for specimen classification with the statistical method of multiple logistic discrimination analysis. The selection of best texture characteristics and classification method was performed using specially developed algorithms in FORTRAN 77. The subset of these characteristics provided 87% correct grading of the specimens and the addition of any of the remaining characteristics did not improve significantly the diagnostic ability of the method. This study demonstrated that texture analysis techniques could provide valuable grading decision support to the pathologists, concerning cancer prognosis.

2. METHODOLOGY

2.1 Collection of histopathological data from biopsies

Collection of histopathological data was performed using needle biopsies that were taken from men with elevated serum PSA levels, or abnormal nodules discovered on digital rectal examination. Specimens were first fixed in 10% neutral formalin buffer. Then the samples were embedded in paraffin and were cut at a thickness of 7 μm. Finally these tissue sections were stained with hematoxylin and eosin for microscopic observation.

In the present study, 50 samples from prostate cancer were collected and studied for the four different Gleason grades. The number of samples in each grade is shown in table 1.

Table 1: Distribution of prostate cancer tissue samples among the four different Gleason grades

	Gleason 2	**Gleason 3**	**Gleason 4**	**Gleason 5**
Number of samples	3	29	16	2

2.2 Digital imaging microscopy system

Images of prostate cancer for Gleason grading were acquired using an Olympus BX-50 light epifluorescent microscope. A 20x UPlanFl (Olympus, NA=0.50) objective lens with corrections for spherical and chromatic aberrations, was used for all experiments. The images were collected by a thermoelectrically cooled color CCD camera (Hitachi KP-C571). The images acquired by the CCD camera were digitized by a PCI color frame grabber (Matrox Meteor, Matrox Electronic Systems Ltd.). Image capture, transferring control and digital image processing was performed using Image Pro Plus (Media Cybernetics). The sensitivity of the camera and the intensity of the illumination source were kept constant during experimentation.

2.3 Gray level Co-occurrence methods

The grey level co-occurrence method of texture is concerned with the spatial distribution and dependence among the grey levels in a local area. In brief, each directional GLCM can be derived from the initial image by estimating the pairwise statistics of pixel intensity. Each element (i, j) of the matrix represents an estimate of the probability of the co-occurrence of pixels with gray values i and j at a given separation, described by a displacement, d and an angle θ. The values of these probability functions can be written in matrix form:

$$\phi(d,\theta) = [\, f(i, j / d, \theta]$$

For different values of d and θ, a different GLCM is computed that is dependent on these values. The algorithm yields a non - symmetric matrix, which has the advantage that only angles up to 180 degrees need to be considered. A single co-occurrence matrix can be obtained from each distance d by averaging four co-occurrence matrices for θ= 0, 45, 90, 135 degrees since the probability matrix for the rest of the directions can be computed from these four directions[18, 16]. Concerning displacement d, values other than 1 are rarely used.

The $\phi(d,\theta)$ is a square matrix of size equal to the number of grey levels in the image. It can be transformed to a symmetric one by adding the GLCM to its transpose and dividing each element by two. Also GLCM is normalized by dividing each element of the matrix by the total number of paired occurrences[17].

These methods can be generally described in three steps of feature extraction. In the first step four directional Grey level Co-occurrence Matrices (GLCM) are computed. In the second step, textural features from these matrices are calculated and in the third step, each textural feature is averaged over the four directions. The second and third stages can be

swapped, averaging the directional GLCMs to give a non dimensional matrix, and computing textural features from this non-directional GLCM[18].

In this work, every RGB image was converted to gray level image of 256 grey levels. Image size was 1000x750 pixels. At first, for every image, the GLCMs were computed for the four basic directions θ=0, 45, 90, 135° and for d=1 pixel. The selection of 1 pixel distance was based on the algorithm proposed by Zucker and Terzopoulos, which chooses the spatial distance for the optimal GLCM that maximizes chi-square significance test[17]. This choice is also in accordance with references of Haralick[18] and Huang[19]. Then, the four directional GLCMs were averaged to yield a non directional GLCM. Afterwards the GLCM was normalized and from this matrix the textural features were calculated. The above mentioned calculations were performed using special algorithms developed in Matalab 7.1..

2.4 Texture features extraction form GLCM

Because of the high dimensionality of the GLCM, the individual elements of the matrix are not used directly for texture analysis. Instead a large number of textural features can be derived from this matrix, with those features proposed by Haralick to be the most important[18].

For the current study, the application of the 13 Haralick features in prostate cancer grading was explored. The equations defining these features can be found in reference 18.

These 13 features, which are independent from rotation of the initial image[18], were calculated from the normalized non-directional GLCM using special algorithms developed in Matlab 7.1..

2.5 Feature selection

From the 13 calculated Haralick textural features, which are used for image description, an appropriate subset should be chosen, that provides the better diagnostic – discriminant ability. In the present study, stepwise variable selection technique was used. This method selects the variables one at a time, starting with the best one, until no significant improvement occurs in diagnostic ability. Textural feature selection was performed using special algorithms developed in FORTRAN 77.

2.6 Classification based on texture features

Image classification and grading based on textural features was performed using special algorithms developed in FORTRAN 77. Classification was performed using multiple group logistic discrimination analysis, a multivariate statistical method. This method enables an image to be classified into one grade on the basis of laboratory profile by means of the likelihood ratio. The basis for categorizing the calculated features data was the probability ratio[20].

$$L\left(x_0\right) = \frac{pr\left(x_0 \mid O\right)}{pr\left(x_0 \mid \bar{O}\right)}$$

where $pr(x_0|O)$ is the probability density function of the observed value X at the point x_0 under the presence of the condition O, while $pr(x_0|\bar{O})$ is the probability density function under the absence of condition O.

For normally distributed populations with different means, μ_O and $\mu_{\bar{O}}$, but with common variance σ^2, the previous equation is transformed to the more simple exponential equation

$$L\left(x\right) = \exp\left(a_0 + a_1 \cdot x\right)$$

where coefficients a_0 and a_1 are calculated as follows

$$a_0 = -\frac{1}{2} \cdot \frac{\mu_O^2 - \mu_{\bar{O}}^2}{\sigma^2}$$

and

$$a_1 = \frac{\mu_O - \mu_{\bar{O}}}{\sigma^2} \tag{1}$$

respectively.

In the occasion that data *X* consist a vector of *k* experimental measurements, on normally distributed populations, the initial equation adopts the following form

$$L(\boldsymbol{x}) = \exp\left(a_0 + \boldsymbol{a}^T \cdot \boldsymbol{x}\right) \tag{2}$$

where

$$a_0 = -\frac{1}{2} \cdot \left(\mu_{\boldsymbol{O}} - \mu_{\bar{\boldsymbol{O}}}\right)^T \cdot \Sigma^{-1} \cdot \left(\mu_{\boldsymbol{O}} + \mu_{\bar{\boldsymbol{O}}}\right) \tag{3}$$

and

$$\boldsymbol{a}^T = \left(\mu_{\boldsymbol{O}} - \mu_{\bar{\boldsymbol{O}}}\right)^T \cdot \Sigma^{-1} \tag{4}$$

with $\boldsymbol{\mu_O}$ and $\boldsymbol{\mu_{\bar{O}}}$ represents the means matrices of the k experimental measurements and $\boldsymbol{\Sigma}$ represents the common covariance matrix.

According to the probability ratio as shown in the initial equation, the critical value results for that *x* where *L(x)=1*, or else where $pr(x_0|O) = pr(x_0|\bar{O})$. This value signifies the categorization line of an experimental measurement in one of the two conditions *O* and $\bar{O}$.

The following diagram (figure 1) summarizes the procedure used for Gleason Grading of prostate cancer biopsies.

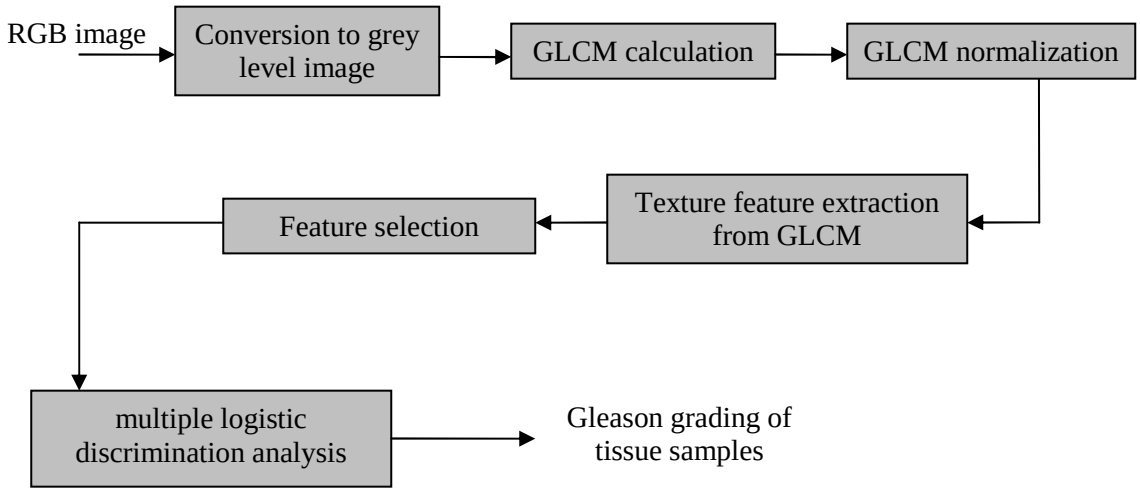

Figure 1: Block diagram of the algorithm used for Gleason Grading of prostate tissue samples

3. RESULTS AND CONCLUSIONS

Representative images of prostate cancer tissue sections are shown in figure 2. In the left coloumn, prostate cancer tissue sections stained with hematoxylin eosin are presented for all Gleason grades (2, 3, 4 and 5). In the right coloumn of figure 2, the corresponding isometric projection of the non-directional GLCM matrix for distance d=1 pixel are shown.

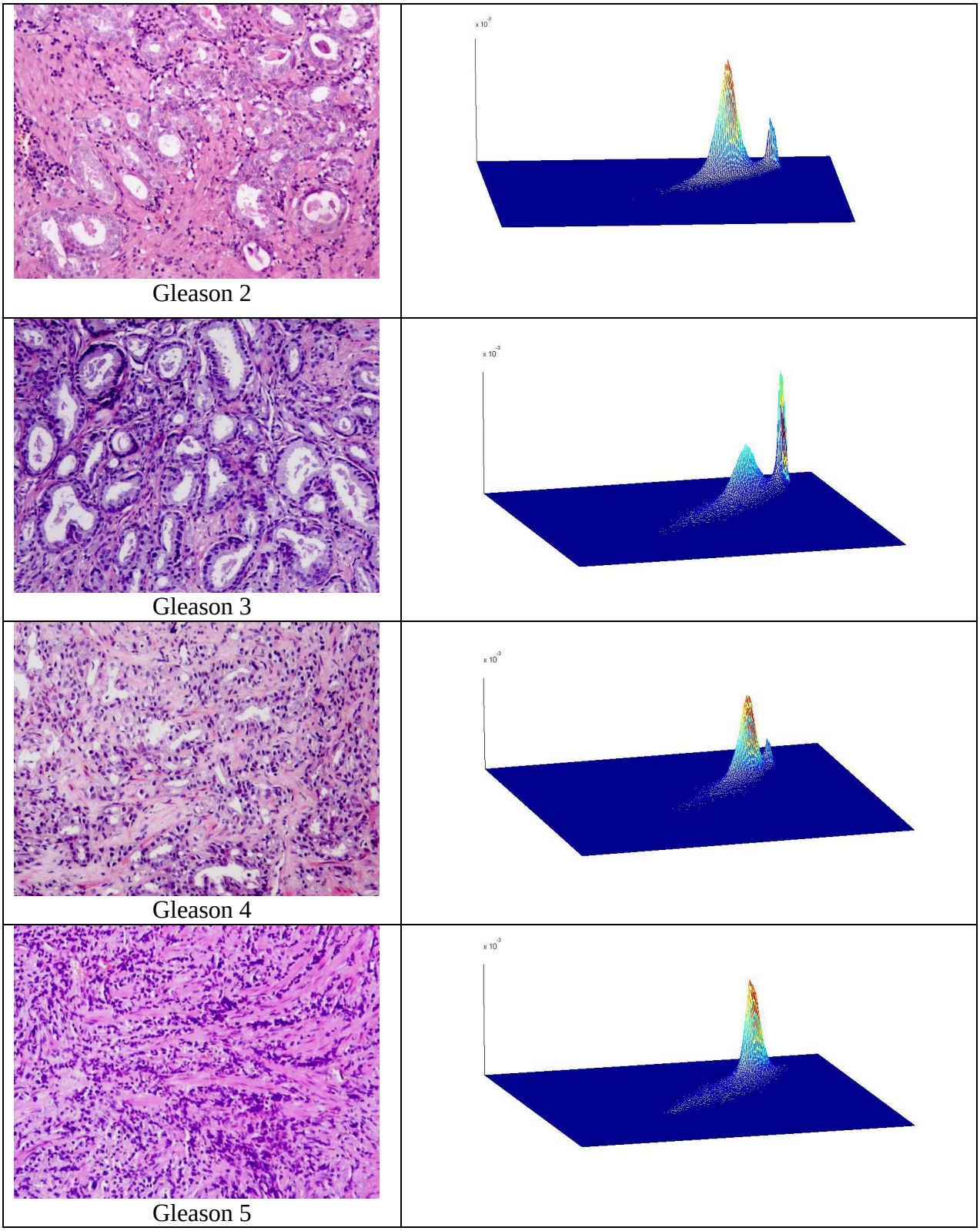

Figure 2: Left column: Representative images of prostate cancer tissue sections stained with hematohylin-eosin and classified to the four Gleason grades. Right column: Corresponding isometric projections of the non-directional GLCM matrix for distance d=1 pixel.

Using stepwise variable selection the features that best discriminate the four Gleason grades were selected. The selected textural features in this case are Sum Entropy, Entropy, Difference Average and Contrast in the order that they appear. In table 2, the mean value and standard deviation are shown for the four textural features as they were calculated from the images for the four different Gleason grades.

Table 2: Mean values of the selected textural features for the four different Gleason grades

Textural features	Gleason 2	Gleason 3	Gleason 4	Gleason 5
Sum Entropy	9.85	10.09	9.40	8.62
Entropy	1.94	2.08	2.29	2.50
Difference average	1.47	2.00	2.27	2.37
Contrast	0.13	0.16	0.18	0.21

Using this subset of textural features in the classification scheme described above, 87% correct classification of the specimens was achieved. The addition of any of the remaining characteristics did not improve significantly the diagnostic ability of the method.

The major problem in computer aided diagnosis or grading of tissue samples, through their microscopic observation, is their description in such a manner that they can be subject of further analysis. Also their significant morphological variation in the different histologic grades should be taken into consideration and enough information should be incorporated to relate each category with its biological background. During classical diagnosis or grading, the histopathologist develops a "mental database" of patterns associated with diseases. Although the clinical utility of this database is unquestionable and remarkable, it is undermined by its subjectivity, experience, lack of quantitation and weakness in transferring to another doctor[11].

In the present study an automated system for prostate cancer biopsy samples classification according to Gleason grading scheme was developed. Such a system brings objectivity and reproducibility to cancer grading and may contribute in prognosis. The developed methodology goes beyond tumor/non-tumor or low/high classification already reported in the literature with a high enough achieved accuracy. Such a system is not going to replace pathologist but instead to provide him with a dynamic tool for more accurate and objective diagnosis, grading and prognosis of the prostate cancer.

4. ACKNOWLEDGMENTS

The project is co - funded by the European Social Fund (75%) and National Resources (25%) - Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly the Program PYTHAGORAS.

REFERENCES

[1] L.A.G. Ries, D. Melbert, M. Krapcho, A. Mariotto, B. A. Miller, E. J. Feuer, L. Clegg, M. J. Horner, N. Howlader, M. P. Eisner, M. Reichman, B. K. Edwards (eds). SEER Cancer Statistics Review, 1975-2004, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2004/, based on November 2006 SEER data submission, posted to the SEER web site, 2007.

[2.] D. F. Gleason, "Classification of prostatic carcinoma", Cancer Chemother. Rep. **50**, 125–128 (1966).

[3] D. G. Bostwick, C. S. Foster, "Predictive factors in prostate cancer: Current concepts from the 1999 College of American Pathologists conference on solid tumor prognostic factors and the 1999 World Health Organization second international consultation on prostate cancer", Sem. Urol. Oncol. **17(4)**, 222-272 (1999).
[4] J. I. Epstein, W. C. Jr Allsbrook, M. B. Amin, L. L. Egevad,. "The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma", Am. J Surg. Pathol. 29, 1228–1242 (2005)
[5] J. Melia, R. Moseley, R. Y. Ball, D. F. R. Griffiths, K. Grigor, P. Harden, M. Jarmulowicz, L. J. McWilliam, R. Montinori, M. Waller, S. Moss, M. C. Parkinson "A UK based investigation of inter- and intra- observer reproducibility of Gleason grading of prostatic biopsies", Histopathology 48(6), 644-654 (2006).
[6] V. Narain, F. J. Bianco, D. J. Grignon, W. A. Sakr, J. E. Pontes, D. P. Wood "How accurately does prostate biopsy Gleason score predict pathologic findings and disease free survival?", The Prostate 49, 185-190 (2001).
[7] M. A. Rubin, R. Dunn, N. Kambham, C. P. Misick, K. M. O'Toole, "Should a Gleason score be assigned to a minute focus of carcinoma on prostate biopsy?", Am. J. Surg. Pathol., 24, 1634-1640 (2000).
[8] M. A. Harding, D. Theodorescu, "Prognostic markers in localized prostate cancer: from microscopes to molecules", Cancer and Metastasis Reviews, 17, 429-437 (1999).
[9] National Cancer Institute (NCI, USA), "Diagnostic Imaging and guided therapy in prostate cancer: SBIR/STTR Initiative Par.", 99 – 149 (1999).
[10] D. G. Bostwick, C. S. Foster, F. Algaba, R. V. P. Hutter, R. Montironi, F. K. Mostofi Prostate tissue factors. In: Murphy G., Denis L., Khoury S., Partin A., Denis L., eds. Prostate cancer. Second international consultation on prostate cancer. WHO and UICC. Plymouth: Plymbridge Distributors. 162-201 (2000).
[11] O. Saidi, C. Cordon-Cardo and J. Costa, "Technology Insight: will systems pathology replace the pathologist?", Nature Clinical Practice Urology, 4, 39-45 (2007)
[12] J. Diamond, N. Anderson, P. Bartels, R. Montironi, P. Hamilton, "The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia", Hum. Pathol. 35(9), 1121–1131 (2004).
[13] M. Roula, J. Diamond, A. Bouridane, P. Miller, A. Amira, "A multispectral computer vision system for automatic grading of prostatic neoplasia", Proceedings of the IEEE International Symposium on Biomedical Imaging, 193-196 (2002).
[14] Y. Smith, G. Zajicek, M. Werman, G. Pizov, Y. Sherman, "Similarity measurement method for the classification of architecturally differentiated images", Comp. Biomed. Res. 32(1), 1–12 (1999).
[15] R. Stotzka, R. Männer, P. H. Bartels, D. Thompson, "A hybrid neural and statistical classifier system for histopathologic grading of prostate lesions", Anal. Quant. Cytol. Histol. 17(3), 204–218 (1995).
[16] A. H. Mir, M. Hanmandlu S. N. Tandon, "Texture analysis of CT images", IEEE Eng. Med. Biol. 14, 781-786 (1995).
[17] S. W. Zucker, D. Terzopoulos, "Finding structure in co-occurrence matrices for textural analysis", Computer Graphics and Image Processing, 12(3), 286-308 (1980).
[18] R. M. Haralick, K. Shanmugam, and I. Dinstein. "Textural features for image classification", IEEE Trans. on Systems, Man and Cybernetics, 3(6),610 - 621, (1973).
[19] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms",. In G. Medioni, R. Nevatia, D. Huttenlocher, and J. Ponce, editors, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 762 - 768 (1997).
[20] A. Albert, E. K. Harris, "Multivariate Interpretation of Clinical Laboratory Data", Marcel Dekker, Inc. 1987)