Cellular Oncology 28 (2006) 85–95
IOS Press


# Prognostic value of adaptive textural features – The effect of standardizing nuclear first-order gray level statistics and mixing information from nuclei having different area

Birgitte Nielsen [a,*] and Håvard E. Danielsen [a,b]
[a] *Department of Medical Informatics, The Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway*
[b] *Division of Genomic Medicine, University of Sheffield, Sheffield S102TN, England*

**Abstract.** *Background*: Nuclear texture analysis is a useful method to obtain quantitative information for use in prognosis of cancer. The first-order gray level statistics of a digitized light microscopic nuclear image may be influenced by variations in the image input conditions. Therefore, we have previously standardized the nuclear gray level mean value and standard deviation. However, there is a clear relation between nuclear DNA content, area, first-order statistics, and texture. For nuclei with approximately the same DNA content, the mean gray level increases with an increasing nuclear area. The aims of the present methodical work were to study: (1) whether the prognostic value of adaptive textural features varies with nuclear area, and (2) the effect of standardizing nuclear first-order statistics. *Methods*: Nuclei from 134 cases of ovarian cancer were grouped into intervals according to nuclear area. Adaptive features were extracted from two different image sets, i.e., standardized and non-standardized nuclear images. *Results*: The prognostic value of adaptive textural features varied strongly with nuclear area. A standardization of the first-order statistics significantly reduced this prognostic information. Several single features discriminated the two classes of cancer with a correct classification rate of 70%. *Conclusion*: Nuclei having an area between 2000–4999 pixels contained most of the class distance information between the good and poor prognosis classes of cancer. By considering the relation between nuclear area and texture, we avoided a loss of information caused by standardizing the first-order statistics and mixing data from cells having different nuclear area.

Keywords: Nuclear texture analysis, early ovarian cancer, prognostic classification, nuclear area, DNA content, nuclear first-order gray level statistics, nuclear area dependent class distance matrices, adaptive features

## 1. Introduction

Nuclear texture analysis is a useful method to obtain quantitative information for the diagnosis and prognosis of human cancer [6,9,12,14,15,26], and gives information about the spatial arrangement of the pixel gray levels in a digitized microscopic nuclear image. Statistical texture methods extract local information from the pixels of the image, and describe the distribution of this information in a statistical way. The extracted statistics may range from simple first-order statistics, such as mean value and standard deviation of the gray level distribution, to second-order or higher order statistics depending on the number of pixels which define the local information.

The first-order gray level statistics of an image are affected by the image input conditions, e.g., the first-order statistics of a digitized light microscopic nuclear image may be influenced by variations in staining, illumination or variations in the photographic process. Therefore, in texture analysis, the images are usually standardized to have the same mean value and standard deviation [24]. It is common practice to perform a histogram matching, i.e., to transform the gray levels of an image so that the normalized histogram of the transformed image matches a specified distribution. Histogram equalization is an example where a match with the uniform distribution is performed. A match of

*Corresponding author: Dr. B. Nielsen, Department of Medical Informatics, The Norwegian Radium Hospital, Montebello, N-0310 Oslo, Norway. Tel.: +47 22 93 48 62; Fax: +47 22 93 56 27; E-mail: birgitn@ifi.uio.no.

the normalized histogram to a Gaussian distribution is more gentle, especially for stochastic textures [4]. Previously, we have performed a linear gray level mapping to standardize the nuclear mean value and standard deviation without standardizing other first-order statistical measures such as kurtosis and skewness [2, 16–19,22,25].

However, in monolayer nuclei, there is a clear relation between nuclear integrated optical density (IOD), nuclear area and nuclear first-order gray level statistics. The optical density (OD) of a pixel is inversely proportional to the pixel gray level value. The IOD, which is a measure of the relative nuclear DNA content, is given by a summation over the nuclear OD pixel values. Therefore, for nuclei with approximately the same DNA content, the mean gray level increases with an increasing nuclear area. For nuclei with approximately the same nuclear area, the mean gray level decreases with an increasing nuclear DNA content. In both cases, we have also observed that the standard deviation of the nuclear gray level distribution decreases with an increasing mean gray level value. These first-order statistical measures influence the nuclear texture (as described by, e.g., the gray level co-occurrence matrix), i.e., the nuclear texture varies with DNA content and nuclear area.

In many popular texture analysis methods, second or higher order statistics on the relation between image gray level values in pixel pairs or sequences of pixels are stored in matrices, e.g., the gray level co-occurrence matrix [10] or the gray level run length matrix [8]. A relative large number of predefined, non-adaptive textural features are then extracted from these matrices. In a unified approach to statistical texture feature extraction, we have proposed to extract only two adaptive features from the matrix of a given texture method [16,18]. The adaptive feature extraction, which is based on class distance matrices and class difference matrices, extracts textural features from the parts of the matrices which actually contain class distance information. In [16], we found that the new adaptive features outperformed the classical features when applied to the most difficult set of 45 Brodatz texture pairs [3]. We also found that one single adaptive feature contained most of the discriminatory power of each of the texture analysis methods.

In order to study the prognostic value of quantifying the chromatin structure of cell nuclei from patients with early ovarian cancer, we have previously extracted nuclear adaptive textural features from pilot material consisting of 40 cases (patients) [16,18]. Class distance and class difference matrices illustrated a clear difference in chromatin texture between the "good" and "poor" prognosis classes.

In the present study, we have extracted class distance matrices and corresponding adaptive features from a series of 134 cases of early ovarian cancer. The monolayer nuclei were grouped into area intervals according to the number of pixels in the nuclei. Separate class distance matrices were extracted from the nuclei within each area interval. Adaptive features based on these area dependent matrices were extracted. Texture analysis was performed on two different image sets, i.e., nuclear images before and after standardization of nuclear first-order gray level statistics.

The aims of the study were: (1) to study whether the prognostic value of adaptive nuclear textural features varies with nuclear area, and (2) to study the effect of standardizing the nuclear first-order gray level statistics.

## 2. Materials and methods

### *2.1. Cell nuclei from early ovarian cancer*

This retrospective study was performed on tissue samples from patients treated for early ovarian cancer during 1982–1989. 134 cases of ovarian cancer classified as International Federation of Gynecology and Obstetrics (FIGO) stage I were included in the analysis. 94 cases had a good prognosis, which means that they survived the follow-up period without a relapse. The minimum length of follow-up for patients alive without a relapse was 10 years. The 40 cases included in the poor prognosis class relapsed or died of a cancer-related disease during the follow-up period.

Paraffin-embedded tissue fixed in 4% buffered formalin was used for preparation of nuclei suspension. The tumor tissue was selected by a pathologist [13]. Monolayers (isolated nuclei) were prepared from one or more 50 $\mu$m sections using a modification of Hedley's method [11]. The nuclei were Feulgen–Schiff stained according to an established protocol [23].

The Fairfield DNA Ploidy System (Fairfield Imaging Ltd, Nottingham, UK), which consisted of a Zeiss Axioplan microscope equipped with a 40/0.75 objective lens (Zeiss), a 546 nm green filter and a black and white high-resolution digital camera (C4742-95, Hamamatsu Photonics K.K., Hamamatsu, Japan) with $1024 \times 1024$ pixels and a gray level resolution of 10 bits/pixel was used to capture each image field. Shad-

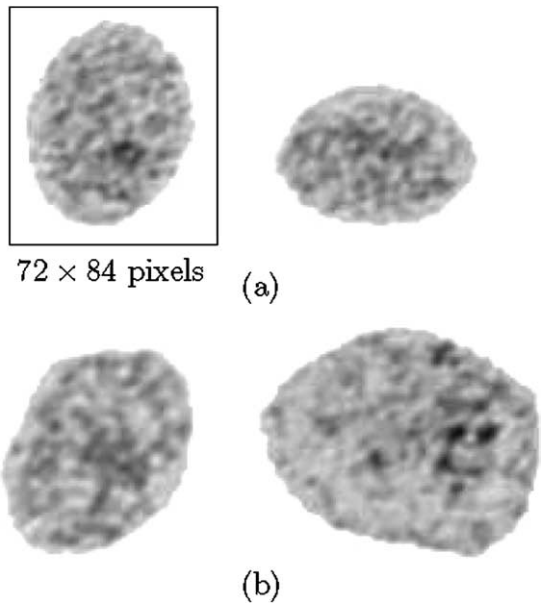


Fig. 1. Isolated cell nuclei (monolayers) from (a) a good prognosis case and (b) a poor prognosis case.

ing correction was performed for each image field. The pixel resolution was 166 nm/pixel on the cell specimen. Trained personnel performed a screening of the nuclei at the microscope and selected tumor nuclei for the analysis. Stromal nuclei, necrotic nuclei, doublets or cut nuclei were disregarded. Each nucleus was segmented from the background by using a global threshold. The segmented nuclei were stored in galleries in each case. The mean number of measured tumor nuclei/case was 281, ranging from 220 to 314 nuclei. Figure 1 shows examples of isolated nuclei from each of the two groups.

### 2.2. Nuclear gray level mean value and standard deviation adjustment

Texture analysis was performed on two different image sets:

#### 2.2.1. Pre-specified nuclear mean value and standard deviation

Each nuclear image was adjusted by a linear pixel value transform to a pre-specified mean value ($\mu_0 = 650.0$) and standard deviation ($\sigma_0 = 120.0$) as in [16, 18]. For every pixel value $i$ in a nuclear image having an original mean value $\mu$ and standard deviation $\sigma$, the linear pixel value mapping

$$v = \mu_0 + \frac{\sigma_0}{\sigma}(i - \mu) \quad (1)$$

gives the new pixel values $v$ of an image having the desired mean value $\mu_0$ and standard deviation $\sigma_0$.

#### 2.2.2. No nuclear mean value and standard deviation adjustment

Here we used the original nuclear images, i.e., the nuclear first-order gray level statistics were not adjusted.

### 2.3. Peel-off scanning

In order to extract separate estimates of texture features in the periphery and center of the nucleus, the 2D gray level nuclear image was transformed into a 1D gray level signal by scanning the nucleus in a spiral-like fashion called "peel-off scanning" [16–18]. The 1D gray level signal was then divided into a peripheral (representing 30% of the total area of the nucleus) and a central (representing 70% of the area) segment. Class distance and class difference matrices and adaptive features were extracted separately from the peripheral (p) and central (c) segments.

### 2.4. Texture analysis from class distance and class difference matrices

Our adaptive feature extraction consisted of the following two steps: (1) Class distance matrices and class difference matrices between the good and poor prognosis classes of cancer were computed from the training set nuclei (see 2.5 for a description of the training and test sets). (2) Based on the computed class distance and difference matrices, adaptive textural features were extracted from all the nuclear images, i.e., nuclei from the training and test sets. Only the training set cases were included in the computation of the class distance and difference matrices, in order to test whether the matrices contained relevant information on the difference between the two classes of cancer. That is, to test whether the matrices were overfitted to the particular training cases.

Class distance matrices and class difference matrices (see [16,18] or the following detailed description) were computed from: Gray Level Cooccurrence Matrices (GLCM) [10], Gray Level Run Length Matrices (GLRLM) [8], Cooccurrence of Gray Level Run Length Matrices (CGLRLM) [1] and Complexity Curves (CC) [2]. The number of gray levels in the images was reduced by re-quantization to $G = 16$ prior to the computation of the GLCM, GLRLM and CGLRLM matrices. The GLCM matrices were computed with an interpixel distance $d = 1$. In the

CGLRLM method we work with two matrices, sum (s) and difference (d) CGLRLM matrices [1]. The complexity curve is one dimensional, and hence we actually compute class distance vectors and class difference vectors. The following Eqs (2)–(7) are valid when we work with matrices. In Section 2.4.2, the one dimensional gray level distribution is used as an example to illustrate the computation of class distance and difference vectors.

#### 2.4.1. Nuclear area independent class distance matrices

For a given texture analysis method, we compute the probability matrices $P_n(i,j \mid \omega_c)$ (e.g. GLCMs), $n = 1,2,\ldots,N(\omega_c)$, were $N(\omega_c)$ is the number of training set nuclear images of class $\omega_c$, $c = 1,2$ ($\omega_1 =$ good prognosis, $\omega_2 =$ poor prognosis). For each element $(i,j)$ in the matrix we then estimate the class conditional probability distribution of the normalized matrix value. Based on these class conditional distributions, we compute the average matrix $\overline{P}(i,j \mid \omega_c)$ for each class $\omega_c$, the class variance matrix $\sigma_P^2(i,j \mid \omega_c)$, the class difference matrix $\Delta(i,j \mid \omega_1,\omega_2)$, and finally the Mahalanobis class distance matrix $J(i,j \mid \omega_1,\omega_2)$ between the two classes $\omega_1$ and $\omega_2$ [16] as follows

$$\overline{P}(i,j \mid \omega_c) = \frac{1}{N(\omega_c)} \sum_{n=1}^{N(\omega_c)} P_n(i,j \mid \omega_c), \tag{2}$$

$$\sigma_P^2(i,j \mid \omega_c) = \frac{1}{N(\omega_c)} \sum_{n=1}^{N(\omega_c)} \big(P_n(i,j \mid \omega_c) - \overline{P}(i,j \mid \omega_c)\big)^2, \tag{3}$$

$$\Delta(i,j \mid \omega_1,\omega_2) = \overline{P}(i,j \mid \omega_1) - \overline{P}(i,j \mid \omega_2), \tag{4}$$

$$J(i,j \mid \omega_1,\omega_2) = 2\frac{(\overline{P}(i,j \mid \omega_1) - \overline{P}(i,j \mid \omega_2))^2}{\sigma_P^2(i,j \mid \omega_1) + \sigma_P^2(i,j \mid \omega_2)}. \tag{5}$$

*Adaptive feature extraction*

Based on the Mahalanobis class distance matrix and the class difference matrix, we extract the following two adaptive features from the probability matrix of a given texture method. Thus, an image $I_k$ with a probability matrix $P_k(i,j)$, $k \in \{1,2,\ldots,K(\omega_c)\}$, were $K(\omega_c)$ is the total number of nuclear images of class $\omega_c$, will give two adaptive feature values [16]

$$F_+ = \sum_{i,j;\Delta(i,j\mid\omega_1,\omega_2)>0} P_k(i,j \mid \omega_c)\big[J(i,j \mid \omega_1,\omega_2)\big]^2,$$

$$F_- = \sum_{i,j;\Delta(i,j\mid\omega_1,\omega_2)<0} P_k(i,j \mid \omega_c)\big[J(i,j \mid \omega_1,\omega_2)\big]^2. \tag{6}$$

#### 2.4.2. Nuclear area dependent class distance matrices

The nuclear images were grouped into area intervals according to the number of pixels in the nucleus ($A_0$: $<1000$ pixels, $A_1$: 1000–1999 pixels, $A_2$: 2000–2999 pixels, …, $A_{10}$: $>10000$ pixels). The number of nuclei within each area interval (in %) was $A_0$: 0.1, $A_1$: 8.5, $A_2$: 24.0, $A_3$: 29.7, $A_4$: 18.7, $A_5$: 9.5, $A_6$: 5.0, $A_7$: 2.2, $A_8$: 1.0, $A_9$: 0.5, $A_{10}$: 0.9. Separate class distance and class difference matrices $J_{A_a}$, $\Delta_{A_a}$ were computed from all training set nuclei within each $A_a$, $a = 1,2,\ldots,10$.

In order to study the difference in first-order gray level statistics between the two classes, class distance and difference vectors were also extracted from the nuclear gray level distributions. Figure 2 illustrates the computation of gray level distribution class distance and difference vectors. The vectors were computed from all nuclei (within nuclear area interval $A_4$) within the good $\omega_1$ and poor $\omega_2$ prognosis classes. First, the gray level distribution $P(i)$ of each nucleus was computed. Then, the average gray level distribution of the good $\overline{P}_{A_4}(i \mid \omega_1)$ and the poor $\overline{P}_{A_4}(i \mid \omega_2)$ prognosis classes (Fig. 2(a)) and the class variance vectors $\sigma^2_{P_{A_4}}(i \mid \omega_1)$ and $\sigma^2_{P_{A_4}}(i \mid \omega_2)$ (Fig. 2(b)) were computed. Finally, the gray level class difference vector $\Delta_{A_4}(i|\omega_1,\omega_2)$ and the gray level Mahalanobis class distance vector $J_{A_4}(i|\omega_1,\omega_2)$ were computed (Fig. 2(c)).

*Adaptive feature extraction*

Adaptive features were computed by selecting the matrices $\Delta_{A_a}$ and $J_{A_a}$ according to the nuclear area of each nucleus

$$F_+ = \sum_{i,j;\Delta_{A_a}(i,j\mid\omega_1,\omega_2)>0} P_k(i,j \mid \omega_c)\big[J_{A_a}(i,j \mid \omega_1,\omega_2)\big]^2,$$

$$F_- = \sum_{i,j;\Delta_{A_a}(i,j\mid\omega_1,\omega_2)<0} P_k(i,j \mid \omega_c)\big[J_{A_a}(i,j \mid \omega_1,\omega_2)\big]^2. \tag{7}$$

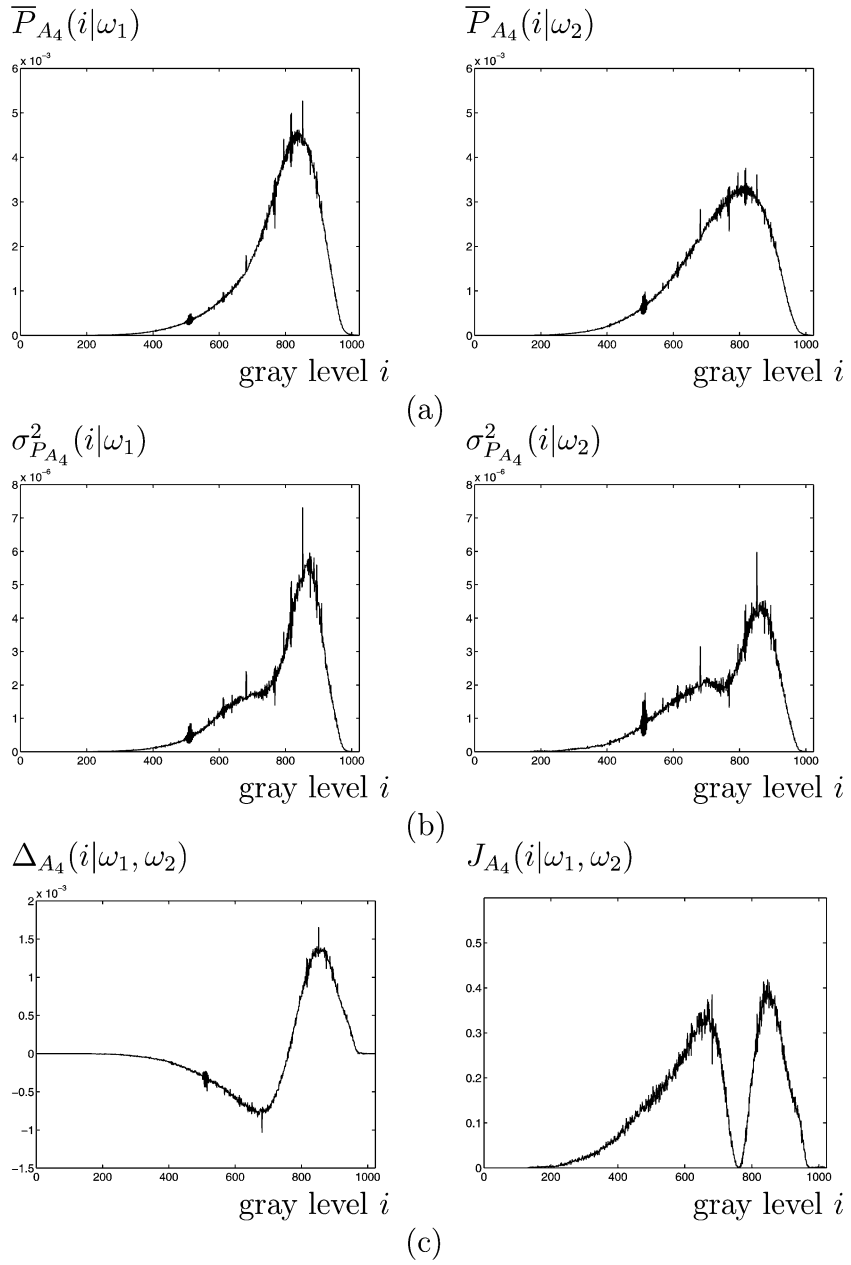


Fig. 2. (a) The average gray level distribution of all nuclear images (within area interval $A_4$, 4000 $\leqslant$ nuclear area $<$ 5000 pixels) of the good $\overline{P}_{A_4}(i \mid \omega_1)$ and the poor $\overline{P}_{A_4}(i \mid \omega_2)$ prognosis classes of early ovarian cancer. (b) The corresponding class variances of the gray level distributions, $\sigma^2_{P_{A_4}}(i \mid \omega_1)$ and $\sigma^2_{P_{A_4}}(i \mid \omega_2)$. (c) The gray level distribution class difference vector $\Delta_{A_4}(i \mid \omega_1,\omega_2) = \overline{P}_{A_4}(i \mid \omega_1) - \overline{P}_{A_4}(i \mid \omega_2)$ and the level gray distribution Mahalanobis class distance vector $J_{A_4}(i \mid \omega_1,\omega_2)$ between the two prognostic classes of cancer. The gray level distributions were computed from the complete non-standardized nuclei.

### *2.5. Classification of cases*

To evaluate the prognostic value of each textural feature, minimum Euclidean distance classifiers [7] based on single features were constructed to classify each case (patient) into the good or poor prognosis classes. In the texture analysis based on area independent class distance matrices, each case was represented by the mean value of the texture feature values extracted from about 280 nuclei/case. In the texture analysis based

on area dependent class distance matrices, the (scalar) mean feature value of each case was computed from the nuclei within area intervals $A_a$, $a = 2, 3, 4$ (about 200 nuclei).

The cases of the two prognosis classes of ovarian cancer were randomly divided into two data sets D1 and D2, each set containing 47 (out of 94) good prognosis and 20 (out of 40) poor prognosis cases. In the first experiment, D1 was used as a training set and D2 as an independent test set. Class distance and difference matrices were first computed from the training set images in D1. Based on these matrices, adaptive features were then extracted from the nuclei in D1. Classifiers based on single features were constructed. In the second experiment, D2 was used as a training set and D1 as a test set. New class distance and difference matrices were computed, adaptive features were extracted, and classifiers were constructed, based on the second training set D2.

The mean training set correct classification rate (CCR) of each feature was computed from the two CCRs obtained from the training sets D1 and D2. For each texture method, the best classifiers (with highest mean training set CCR) were tested on the test sets D2 (experiment 1) and D1 (experiment 2). From these two experiments, the mean test set CCR was computed.

## 3. Results

### 3.1. Class distance and difference matrices

Figure 3 shows the distance and difference in the nuclear gray level distribution between the two prognostic classes of cancer. With increasing nuclear area the higher probabilities of the GLCM- and GLRLM-matrices were shifted towards the locations of the matrices corresponding to lighter gray levels. As an example of this, compare the average GLCM-matrices computed from area intervals $A_1$ and $A_4$ (Fig. 4).

The elemental values of the class distance and class difference matrices (vectors) varied with $A_a$, $a = 1, 2, \ldots, 10$ (see Figs 3, 5 and 6). For each texture method the class distance matrices computed from area intervals $A_3$ and $A_4$ contained most of the prognostic information between the two classes of cancer. The Mahalanobis values contained in $J_{A_4}(i, j)$ were about five times higher than the Mahalanobis values contained in the corresponding area independent matrices $J(i, j)$. By standardizing nuclear first-order statistics the class distances contained in $J_{A_3}(i, j)$ and $J_{A_4}(i, j)$ were decreased (compare Figs 5 and 7).

Matrices computed from $A_1$ contained almost no class distance information (see Figs 3, 5, 6). With increasing nuclear area, the area dependent class distance and difference matrices were based on a decreasing number of nuclei, resulting in more noise in the matrices (see Fig. 3).

### 3.2. Adaptive features

Adaptive features based on nuclear area dependent matrices were extracted from all nuclei within area intervals $A_a$, $a = 2, 3, 4$ (which contained 72.4% of the nuclei). Nuclei within area intervals $A_1$ and $A_a$, $4 < a \leqslant 10$, were excluded from the analysis (see the last paragraph of 3.1).

For each of the four combinations (area independent matrices, standardized nuclei), (area independent matrices, non-standardized nuclei), (area dependent matrices, standardized nuclei), (area dependent matrices, non-standardized nuclei), the best single adaptive feature (i.e., the feature with the highest mean training set correct classification rate, see 2.5) of each of the four texture methods, were evaluated on the test sets. The mean training and test set correct classification rates of the best single features are given in Table 1. Each of the three combinations (area independent matrices, standardized nuclei), (area independent matrices, non-standardized nuclei), and (area dependent matrices, standardized nuclei) gave a correct classification rate of about 60%, while the fourth combination (area dependent matrices, non-standardized nuclei) discriminated the two prognostic classes of cancer with a correct classification rate of about 70%.

## 4. Discussion

Nuclear first-order gray level statistics may be influenced by variation in staining, illumination or other variations in the monolayer preparation and imaging process. Therefore, in previous studies on nuclear texture analysis, we have standardized the first-order statistics. In [16], the mean feature values extracted from all nuclei/case gave a 60–65% correct classification rate, regardless of the texture analysis method (GLCM, GLRLM, CGLRLM, CC). These results were based on pilot material consisting of 40 cases. In the present study, we have verified that these features give a correct classification rate of 60–65%, see Table 1 (area in-

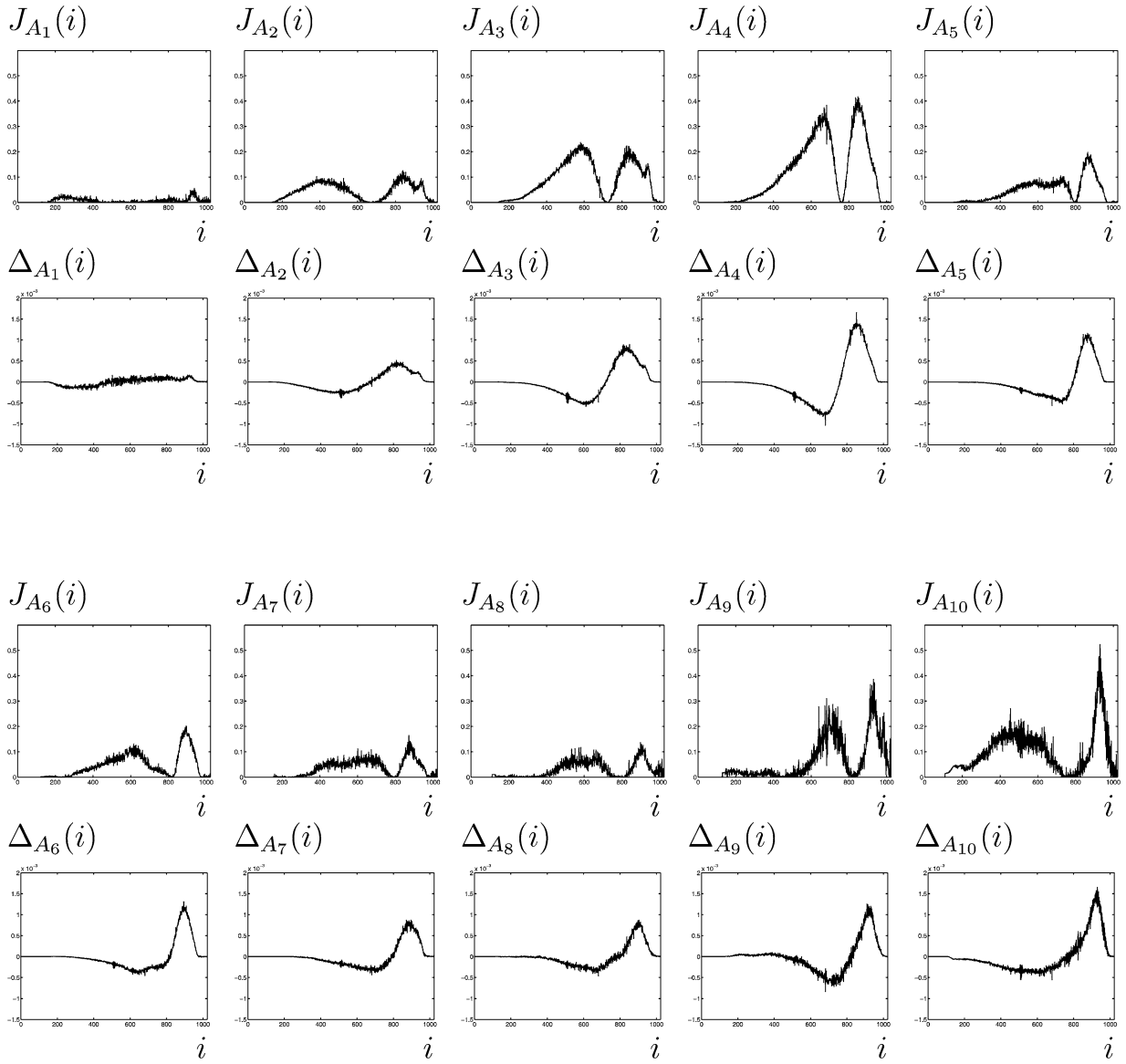


Fig. 3. The gray level distribution Mahalanobis class distance vector $J_{A_a}(i \mid \omega_1, \omega_2)$ and the gray level distribution class difference vector $\Delta_{A_a}(i \mid \omega_1, \omega_2)$ between the good $\omega_1$ and poor $\omega_2$ prognosis classes of early ovarian cancer, computed from nuclei within each area interval $A_a$, $a = 1, \ldots, 10$. The positive (negative) values in the difference vectors correspond to gray levels $i$ that are more (less) probable for the good prognosis class compared to the poor prognosis class. The vectors were computed from the complete non-standardized nuclei.

dependent, standardized). We have also extracted fractal features from the pilot material [18] and the more complete material, i.e., the same material as used in the present study [19]. In [19], the fractal dimension feature gave a correct classification rate of 62.7% when evaluated on the test set. As in the present study, we found that the area intervals $A_3$ and $A_4$ contained most of the class distance information. The best combination of two lacunarity features computed from these area intervals gave a correct classification rate of 62.7% on the test set (compare with the results in Table 1 (area dependent, standardized)).

In the present study, we have shown that when working with area dependent matrices, a standardization of the nuclear first-order statistics significantly reduces the class distance. Single GLCM, GLRLM and CGLRLM textural features, extracted from non-standardized nuclei within area intervals $A_a$, $a = 2, 3, 4$, discriminated the two prognostic classes with a correct classification rate of about 70%, see Table 1

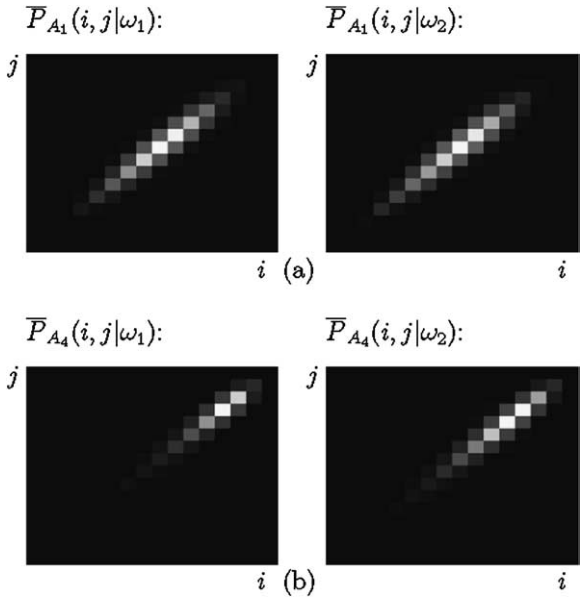


Fig. 4. The average Gray Level Cooccurrence Matrix (GLCM) $\overline{P}_{A_a}(i,j \mid \omega_c)$ of the good prognosis $\omega_1$ and the poor prognosis $\omega_2$ classes of early ovarian cancer, computed from (a) all nuclei within area interval $A_1$ ($1000 \leqslant$ area $< 2000$ pixels) and (b) all nuclei within area interval $A_4$ ($4000 \leqslant$ area $< 5000$ pixels). Lighter elements in the gray-scale plots correspond to higher probabilities. The matrices were computed from the 70% central part of the non-standardized nuclei.

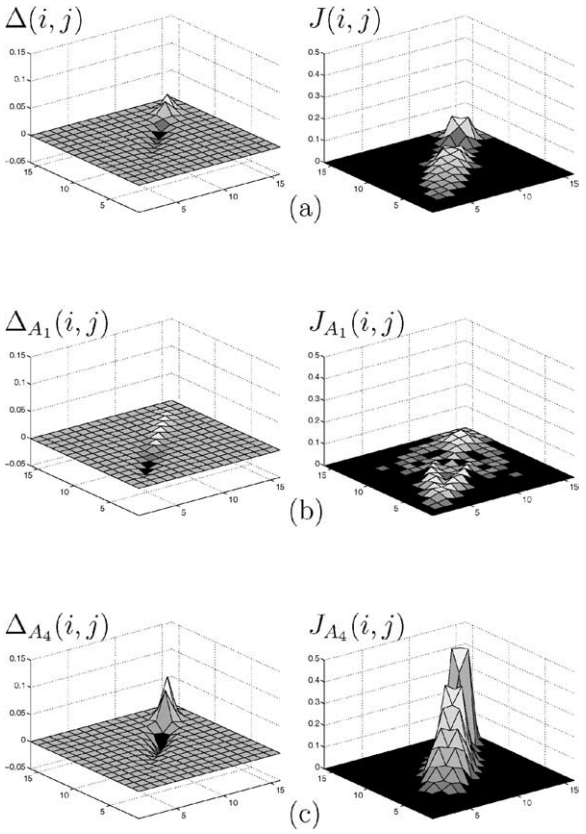


Fig. 5. Gray Level Cooccurrence Matrix (GLCM) Mahalanobis class distance matrices $J(i,j)$ and GLCM class difference matrices $\Delta(i,j)$ between the two prognostic classes of early ovarian cancer, computed from nuclear images with non-standardized first-order gray level statistics. The matrices were computed from (a) all 37,690 nuclei, (b) all nuclei within area interval $A_1$ ($1000 \leqslant$ nuclear area $< 2000$ pixels) and (c) all nuclei within area interval $A_4$ ($4000 \leqslant$ nuclear area $< 5000$ pixels). The positive (negative) values in the difference matrices correspond to GLCM matrix elements that are more (less) probable for the good prognosis class compared to the poor prognosis class. The matrices were computed from the 70% central part of the nuclei.

(area dependent, non-standardized). There was a high correlation between these adaptive features. In experiment 1 (see 2.5), 38 of the misclassified patients (the number of misclassified patients was; GLCM feature $F_{-(c)}$: 39, GLRLM feature $F_{-(c)}$: 40, CGLRLM feature $F_{s-(c)}$: 40) were misclassified by each of these texture features. In experiment 2, 37 of the misclassified patients (the number of misclassified patients was; GLCM feature $F_{-(c)}$: 39, GLRLM feature $F_{-(c)}$: 40, CGLRLM feature $F_{s-(c)}$: 39) were misclassified by each of these features. Almost all of the misclassified patients were misclassified in both experiments. The nearly identical classification results of experiments 1 and 2, which included different cases in the training sets (see 2.5), verified that the class distance and difference matrices and the classifiers were not overfitted to the particular training cases.

There is a complex relation between nuclear size, DNA content, gray level distribution and texture. Even diploid monolayer nuclei from a given patient vary greatly in size, in first order gray level moments as well as in nuclear texture. In the present study, the nuclear area of the monolayer nuclei varied between 1000 and 30,000 pixels. Our adaptive texture feature extraction is based on large areas of consistently high values within the class distance matrices. When working with original nuclear images (i.e., non-standardized images), we have shown that with increasing nuclear area (area intervals $A_a$, $a = 1, 2, 3, \ldots, 10$) the locations of the most class discriminating matrix elements are shifted towards locations of the matrix corresponding to lighter gray levels. Therefore, it is important to avoid mixing data from different area intervals in the accumulation of the class distance and difference matrices.

On the other hand, because of this relation between nuclear area, first-order statistics, and the location of

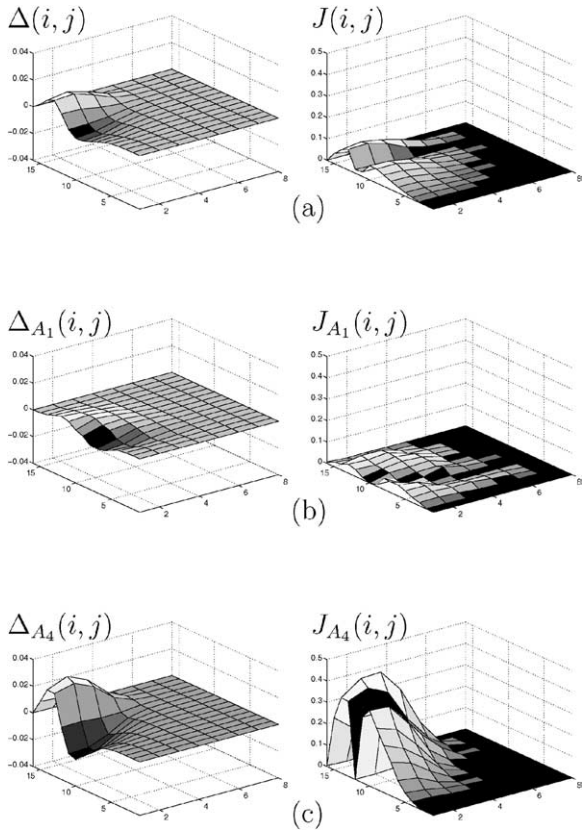


Fig. 6. Gray Level Run Length Matrix (GLRLM) Mahalanobis class distance matrices $J(i,j)$ and GLRLM class difference matrices $\Delta(i,j)$ between the two prognostic classes of early ovarian cancer, computed from nuclear images with non-standardized first-order gray level statistics. The matrices were computed from (a) all 37,690 nuclei, (b) all nuclei within area interval $A_1$ ($1000 \leqslant$ nuclear area $< 2000$ pixels) and (c) all nuclei within area interval $A_4$ ($4000 \leqslant$ nuclear area $< 5000$ pixels). The positive (negative) values in the difference matrices correspond to GLRLM matrix elements that are more (less) probable for the good prognosis class compared to the poor prognosis class. The matrices were computed from the 70% central part of the nuclei.

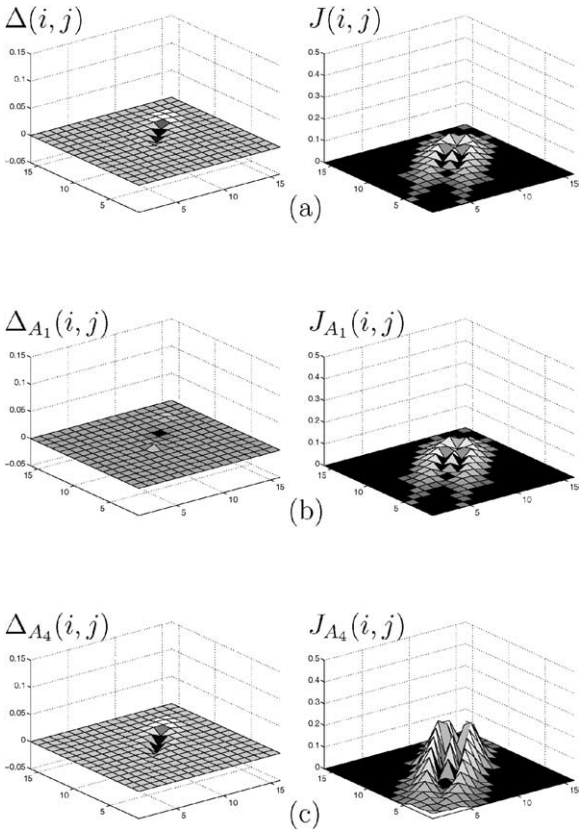


Fig. 7. Gray Level Cooccurrence Matrix (GLCM) Mahalanobis class distance matrices $J(i,j)$ and GLCM class difference matrices $\Delta(i,j)$ between the two prognostic classes of early ovarian cancer, computed from nuclear images with standardized first-order gray level statistics. The matrices were computed from (a) all 37,690 nuclei, (b) all nuclei within area interval $A_1$ ($1000 \leqslant$ nuclear area $< 2000$ pixels) and (c) all nuclei within area interval $A_4$ ($4000 \leqslant$ nuclear area $< 5000$ pixels). The positive (negative) values in the difference matrices correspond to GLCM matrix elements that are more (less) probable for the good prognosis class compared to the poor prognosis class. The matrices were computed from the 70% central part of the nuclei.

the most class discriminating matrix elements, it is important to standardize first-order statistics when accumulating area independent matrices. From the standardized images, we obtained similar results from features based on area independent and area dependent matrices.

We have shown that class distance and difference values vary greatly with nuclear area interval $A_a$, $a = 1, 2, \ldots, 10$. Class distance and difference matrices computed from area intervals $A_a$, $4 < a \leqslant 10$ (corresponding to nuclear area between 5000–30,000 pixels) were based on a decreasing number of nuclei, resulting in more noise in the matrices. Therefore, nuclei within these area intervals were excluded from the analysis. However, these large nuclei may contain prognostic information. If a larger number of nuclei were measured for each patient, some of these matrices could become more stable and larger nuclei could be included in the analysis.

We have previously proposed a unified approach to low dimensional adaptive texture feature extraction. Through several studies we have found that, for a given texture analysis method, one single adaptive feature contains most of the discriminatory power of the method. In the present methodical study, we have improved the performance of these adaptive features

Table 1

The mean training (train) and test set correct classification rates, CCR (in %), of the best single adaptive features (given in parenthesis) of each of the texture analysis methods

| Method | Area independent | | Area dependent | |
|---|---|---|---|---|
| | Standardized | Non-standardized | Standardized | Non-standardized |
| | CCR train, CCR test | CCR train, CCR test | CCR train, CCR test | CCR train, CCR test |
| GLCM | 63.4, 62.7 ($F_{-(p)}$) | 57.5, 57.5 ($F_{-(c)}$) | 67.9, 62.7 ($F_{-(p)}$) | 70.9, 70.9 ($F_{-(c)}$) |
| GLRLM | 65.7, 64.2 ($F_{+(c)}$) | 60.5, 59.0 ($F_{+(c)}$) | 67.2, 61.2 ($F_{-(p)}$) | 70.9, 69.4 ($F_{-(c)}$) |
| CGLRLM | 66.4, 61.9 ($F_{s-(p)}$) | 59.7, 59.0 ($F_{s+(c)}$) | 68.7, 61.9 ($F_{s-(c)}$) | 70.9, 70.2 ($F_{s-(c)}$) |
| CC | 60.5, 61.2 ($F_{+(p)}$) | 62.7, 59.0 ($F_{+(p)}$) | 66.4, 59.7 ($F_{-(p)}$) | 70.9, 66.4 ($F_{-(p)}$) |

The features were computed from area independent and area dependent class distance and difference matrices, and for two different data sets; standardized images, i.e., nuclear images with standardized 1. order statistics and non-standardized images, i.e., original nuclear images. GLCM: Gray Level Cooccurrence Matrix, GLRLM: Gray Level Run Length Matrix, CGLRLM: Cooccurrence of Gray Level Run Length Matrix, s: sum CGLRLM, CC: Complexity Curve, (p): peripheral part of the nuclei, (c): central part of the nuclei.

when applied to prognostic classification of cancer patients. We now have single texture features that discriminate the two prognostic classes of early ovarian cancer with a correct classification rate of about 70%.

It would be interesting to combine textural features with other methods (prognostic factors), e.g., DNA ploidy analysis, to see whether textural features could add prognostic information. DNA content estimation is performed on extinction images, i.e., optical density images, which are extinction standardized using white value estimation from near background. Several researchers have extracted nuclear texture features from extinction images [5,20,21]. It would also be interesting to compare the results of the present study with features extracted from such extinction standardized optical density images. Nuclear texture analysis and DNA ploidy analysis could then be performed on the same image set.

In conclusion, nuclei having an area between 2000–4999 pixels contained most of the class distance information between the good and poor prognosis classes of early ovarian cancer. By considering the relation between nuclear area, first-order gray level statistics and texture, we avoided a loss of information caused by standardizing the first-order statistics and mixing data from cells having different nuclear area.

## Acknowledgements

The present work was supported by the Research Council of Norway and the Department of Informatics, University of Oslo. The authors would like to thank Prof. F. Albregtsen of the Department of Informatics, University of Oslo, for useful discussions.

## References

[1] F. Albregtsen and B. Nielsen, Texture classification based on co-occurrence of gray level run length matrices, *Australian Journal of Intelligent Information Processing Systems* **6**(1) (2000), 38–45.

[2] S. Baheerathan, F. Albregtsen and H.E. Danielsen, New texture features based on the complexity curve, *Pattern Recognition* **32**(4) (1999), 605–618.

[3] P. Brodatz, *Textures – A Photographic Album for Artists & Designers*, Dover, New York, 1966.

[4] J.M. Carstensen, Description and simulation of visual texture, The Technical University of Denmark, Technical Report PhD, no. 59, 1992.

[5] H. Christen, M. Oberholzer, M. Buser, R. Lötscher, R. Gschwind, F. Rösel, R. Ettlin, A. Feess and P. Dalquen, Digital image analysis in cytological diagnosis: a morphometric analysis on pleural mesotheliomas, *Analytical Cellular Pathology* **1** (1989), 105–122.

[6] T. Dreyer, I. Knoblauch, A. Doudkine, C.M. MacAulay, D. Garner, B. Palcic and C. Popella, Nuclear texture features for classifying benign vs. dysplastic or malignant squamous epithelium of the larynx, *Analyt. Quant. Cytol. Histol.* **23** (2001), 193–200.

[7] R.O. Duda, P.E. Hart and D. Stork, *Pattern Classification*, 2nd edn, Wiley-Interscience, New York, 2001.

[8] R.M.M. Galloway, Texture analysis using gray level run lengths, *Computer Graphics and Image Processing* **4** (1975), 172–179.

[9] A. Gschwendtner, Y. Hoffmann-Weltin, G. Mikuz and T. Mairinger, Quantitative assessment of bladder cancer by nuclear texture analysis using automated high resolution image cytometry, *Mod. Pathol.* **12** (1999), 806–813.

[10] R.M. Haralick, K. Shanmugam and I. Dinstein, Textural features for image classification, *IEEE Trans. on Systems, Man and Cybernetics* **3** (1973), 610–621.

[11] D.W. Hedley, DNA analysis from paraffin-embedded blocks, *Methods Cell Biol.* **41** (1994), 231–240.

[12] T. Jørgensen, K. Yogesan, F. Skjørten, O. Kaalhus, K.J. Tveter and H.E. Danielsen, Nuclear texture analysis: A new prognostic tool in metastatic prostate cancer, *Cytometry* **24** (1996), 277–283.

[13] G.B. Kristensen, W. Kildal, V.M. Abeler, J. Kaern, I. Vergote, C.G. Tropé and H.E. Danielsen, Large-scale genomic instability predicts long-term outcome for women with invasive stage I ovarian cancer, *Annals of Oncology* **14** (2003), 1494–1500.

[14] T. Mairinger, G. Mikuz and A. Gschwendtner, Nuclear chromatin texture analysis of nonmalignant tissue can detect adjacent prostatic adenocarcinoma, *The Prostate* **41** (1999), 12–19.

[15] A. Neher, G. Ofner, E. Appenroth and A. Gschwendtner, High-resolution image cytometry on smears of normal oral mucosa: a possible approach for the early detection of laryngopharyngeal cancers, *Head & Neck* **26** (2004), 694–700.

[16] B. Nielsen, F. Albregtsen and H.E. Danielsen, Low dimensional adaptive texture feature vectors from class distance and class difference matrices, *IEEE Transactions on Medical Imaging* **23**(1) (2004), 73–84.

[17] B. Nielsen, F. Albregtsen and H.E. Danielsen, The use of fractal features from the periphery of cell nuclei as a classification tool, *Analytical Cellular Pathology* **19** (1999), 21–37.

[18] B. Nielsen, F. Albregtsen, W. Kildal and H.E. Danielsen, Prognostic classification of early ovarian cancer based on very low dimensionality adaptive texture feature vectors from cell nuclei from monolayers and histological sections, *Analytical Cellular Pathology* **23** (2001), 75–88.

[19] B. Nielsen, F. Albregtsen and H.E. Danielsen, Fractal analysis of monolayer cell nuclei from two different prognostic classes of early ovarian cancer, in: *Fractals in Biology and Medicine*, Vol. 4, Birkhäuser-Verlag, Basel, 2005, pp. 175–186.

[20] K. Rodenacker, Invariance of textural features in image cytometry under variation of size and pixel magnitude, *Analytical Cellular Pathology* **8** (1995), 117–133.

[21] K. Rodenacker and E. Bengtsson, A feature set for cytometry on digitized microscopic images, *Analytical Cellular Pathology* **25** (2003), 1–36.

[22] H. Schulerud, G.B. Kristensen, K. Liestøl, L. Vlatkovic, A. Reith, F. Albregtsen and H.E. Danielsen, A review of caveats in statistical nuclear image analysis, *Analytical Cellular Pathology* **16** (1998), 63–82.

[23] H.J. Tanke and E.M. van Ingen, A reliable Feulgen-acriflavine-$SO_2$ staining procedure for quantitative DNA measurements, *J. Histochem. Cytochem.* **28** (1980), 1007–1013.

[24] F. Tomita and S. Tsuji, *Computer Analysis of Visual Textures*, Kluwer Academic Publishers, Massachusetts, 1990.

[25] K. Yogesan, H. Schulerud, F. Albregtsen and H.E. Danielsen, Ultrastructural texture analysis as a diagnostic tool in mouse liver carcinogenesis, *Ultrastructural Pathology* **22**(1) (1998), 27–37.

[26] K. Yogesan, T. Jørgensen, F. Albregtsen, K.J. Tveter and H.E. Danielsen, Entropy-based texture analysis of chromatin structure in advanced prostate cancer, *Cytometry* **24** (1996), 268–276.