Journal of Korea Multimedia Society Vol. 11, No. 12, DECEMBER 2008(pp. 1635-1648)

Evaluation of Volumetric Texture Features for Computerized Cell Nuclei Grading

Tae-Yun Kim[†], Hyun-Ju Choi[††], Heung-Kook Choi[†††]

ABSTRACT

The extraction of important features in cancer cell image analysis is a key process in grading renal cell carcinoma. In this study, we applied three-dimensional (3D) texture feature extraction methods to cell nuclei images and evaluated the validity of them for computerized cell nuclei grading. Individual images of 2,423 cell nuclei were extracted from 80 renal cell carcinomas (RCCs) using confocal laser scanning microscopy (CLSM). First, we applied the 3D texture mapping method to render the volume of entire tissue sections. Then, we determined the chromatin texture quantitatively by calculating 3D gray-level co-occurrence matrices (3D GLCM) and 3D run length matrices (3D GLRLM). Finally, to demonstrate the suitability of 3D texture features for grading, we performed a discriminant analysis. In addition, we conducted a principal component analysis to obtain optimized texture features. Automatic grading of cell nuclei using 3D texture features had an accuracy of 78.30%. Combining 3D textural and 3D morphological features improved the accuracy to 82.19%. As a comparative study, we also performed a stepwise feature selection. Using the 4 optimized features, we could obtain more improved accuracy of 84.32%. Three dimensional texture features have potential for use as fundamental elements in developing a new nuclear grading system with accurate diagnosis and predicting prognosis.

Key words: Digital Image Cytometry, 3D Texture Analysis, Feature Extraction

1. INTRODUCTION

The evaluation of cell nuclei features helps determine the prognoses of patients with carcinomas, and from a cytometric perspective, the variation in the distribution of chromatin is considered a very important characteristic.

Chromatin is distinguished from heterochromatin and euchromatin by the concentration of DNA. Qualitative identification of the nucleus must be completed using quantitative measurements to establish the state of DNA in the nucleus, which provides an indirect representation of chromatin organization.

Various computer-assisted diagnosis systems have been developed, and in most of these systems, feature extraction is an important means of classification because the classification performance depends on the extracted features. The most critical problems with the available systems are that they still depend for the most part on the pathologist's subjective decision. Conventional visual analysis for grading has low reproducibility because it is based on subjective evaluation, which is prone to

※ Corresponding Author : Heung-Kook Choi, Address : (621-749) School of Computer Science, Inje University, Obang-dong 607, Gimhae, Korea TEL : +82-55-320-3437, FAX : +82-55-322-3107, E-mail : cschk@inje.ac.kr
Receipt date : Oct. 29, 2007, Approval date : June 27, 2008
[†] Department of Computer Science, Graduate School, Inje University
(E-mail : liminus@paran.com)
[††] BK21 Medical Science Education Center, School of Medicine, Busan National University
(E-mail : hjchoi@pusan.ac.kr)
[†††] Ubiquitous Healthcare Research Center(UHRC), Inje University
※ This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2006-311-D 00840) and also supported by the Korea Research Foundation Grant funded by the Korea Government (MOEHRD, Basic Research Promotion Fund) KRF-2005-217-D00 005).

inter- and intra-observer variation, regardless of any grading system.

Several studies have examined and defined various texture features, such as the internal structure of cells (granularity and regularity of chromatin), size irregularity, shape of the nucleus, and distance between nuclei, which are important for determining the progress of cancer[1-6]. Histology-based statistical analyses of textural features are generally based on the gray level of cell nuclei, while the structural analysis method describes the properties and placement of texture elements. Nevertheless, most two-dimensional (2D) texture feature based analysis systems still have low objectivity and reproducibility. Given the variety of analysis methods, no clear measurement standard has been established for extracting accurate numerical information. Similarly, the image analysis systems based on 2D images have several intrinsic limitations. For example, cells and cell nuclei are not perfectly spherical, and consequently, their shape differs noticeably according to the cutting angle and thickness of the sample tissues. Ultimately, it is difficult to confirm the shape of a cell. Another drawback of conventional 2D slice-based approaches is that they are tedious, fatiguing, and time-consuming. To guarantee reproducibility, a new method based on three-dimensional (3D) image analysis is required.

Recently, some papers have reported various 3D texture features. Jafari-Khouzani et al. suggested an analysis method based on a comparative study of 2D and 3D wave let features[7]. Madhabushi et al. studied the automatic segmentation of high-resolution magnetic resonance (MR) images using a 3D Gabor filter and a co-occurrence matrix[8]. Kurani and Xu applied a 3D gray-level co-occurrence matrix (GLCM) and a 3D gray-level run length method (GLRLM) to computed tomography (CT) images to separate various organs of the human body[9-10]. Most of these approaches were simply extended from conventional 2D methods, but the importance of 3D texture increases with its successful expansion.

This study checks the validity of 3D GLCM and 3D GLRLM, which were studied by Kurani and Xu, by applying them to the images of cell nuclei for RCC (Renal Cell Carcinoma) obtained by CLSM. The previous study examined the correlation between the changing grade following the cancer process and the 3D morphological features and also investigated the 2D features that can be a good proxy for estimating 3D features, while mainly focusing on the morphological changes of the nucleus[11-12]. In contrast to the previous study about the external changes of the nucleus, this study focuses on the delicate changes in chromatin pattern inside the nucleus.

In what follows, Section 2 describes the research method along with the details of obtaining the images used in the experiment. Section 3 and Section 4 present the experiment and study results from the application of the method described in Section 2 to the actual image data. Finally, Section 5 contains concluding remarks evaluating the results of the study as well as suggesting a direction for further studies.

2. MATERIAL AND METHODS

2.1 Image Acquisition

We obtained eight classes of RCC tissue from the Department of Pathology, Yonsei University, Korea. They had been fixed in 10% neutral-buffered formalin and embedded in paraffin before receipt. The tissues were cut into 20-μm sections, stained with propidium iodide (PI) containing RNase A at a final concentration 0.5 mg/mL, and mounted in fluorescent mounting medium (DAKO, Carpinteria, CA, USA). The RCC tissues were imaged under a TCS SP2 AOBS confocal laser scanning microscope (Leica Microsystems, Mannheim, Germany), with a 630x, 2x zoom, 1.4 NA HEX PL-Apochromat objective lens, and a HeNe laser.

A series of 2D optical sections, 0.4-μm apart, were acquired, starting above the top surface of the section and extending down to the bottom surface. We obtained 100-130 slices for each volumetric data set, and each slice was a 24-bits/pixel image with a resolution of 256 * 256 pixels.

2.2 Volume Rendering

To extract 3D texture features from volume data and inspect the entire shape of cell nuclei, we implemented a volume rendering module using OpenGL® Shading Language (GLSL)[13-14]. We applied direct volume rendering methods based on 3D texture mapping. This method proved to be very suitable for visualizing the 3D scalar field. The basic concept of the 3D-texture mapping approach is to use the scalar field as a 3D texture. At the core of the algorithm, multiple equidistant planes (slices) parallel to the image plane are clipped against the bounding box of the volume. During rasterization, hardware is used to interpolate 3D-texture coordinates at the polygon vertices and to reconstruct the texture samples using trilinear interpolation within the volume. Finally, the 3D representation is produced by successive blending of the textured polygons back-to-front onto the viewing plane. Since this process uses the blending and interpolation capabilities of the underlying hardware, the time taken to generate an image is negligible compared to software-based approaches. As an advantage of this approach, interactive frame rates are achieved, even if applied to high-resolution scalar fields.

We also used the graphical processor unit-based 3D slicing technique to reduce the computation cost on the CPU. This method consists of two components : a vertex shader, which manages plane objects that should be clipped on the GLSL and a fragment shader, which determines the values of the color buffers to be displayed. Unlike the conventional equation calculated from viewing matrices, vertices, and boundaries that define the texture space, it uses only the rotation transform matrices as input values. Consequently, we can easily reduce computational costs.

2.3 2D and 3D Texture feature extraction

Computer-aided image analysis provides more objective numerical information compared to visual analysis by humans, which is one of the merits of digital image analysis. The 3D texture analysis, which has been commonly used on the clinical data in a number of recent studies, was employed for quantification, which was subsequently analyzed to see whether it can be used as a feature for classification. In order to compare the validity of the extracted features, several types of statistical classifiers were created using 3D texture feature and 2D feature and the classification results among them were subsequently compared.

2.3.1 Spatial relationship in 2D and 3D

The difference of spatial position of two image elements (pixels and voxels) can be described by a displacement vector. In 2D, for a certain distance D, there are 8 neighboring pairs independent directions corresponding to $\phi = 0^\circ, 45^\circ, 90^\circ, 135^\circ$.

In volumetric data, the displacement vectors still can be decomposed into a norm-1 distance D and a direction which can be specified by azimuth ϕ and zenith θ. there are totally 26 neighboring voxel-pairs in 13 independent directions.

As like this, the difference between 2D and 3D data for calculating GLCM and GLRLM lie in the

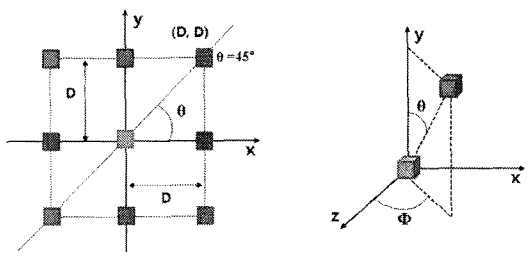


Fig. 1. Spatial relationship in 2D and 3D: (a) 2D (b)3D

displacement vectors.

2.3.2 Feature calculation

Generally, 2D texture features are computed using pixels from each slice. However, if we process 3D volume data as individual 2D slices, some inter-slice information is ignored, increasing the possibility of data loss. To resolve such problems, we applied the concept of 3D texture features. Despite the simplicity of extending conventional matrix-based algorithms to three dimensions, this approach gave a noticeable result. The 2D GLCM considers the spatial dependency of pixels on one slice, while 3D GLCM quantifies the 3D dependency of voxel data on the object volume, which exists across several slices. Similar to the case for

Table 1. The list of GLCM and GLRLM texture feature

Gray Level Co-occurrence Matrix(GLCM)		
Angular Second Moment	$ASM=\sum_i\sum_j p(i,j)^2$	2D, 3D
Contrast	$CONT=\sum_{i-j=0}^{G-1}(i-j)^2\sum_{i=1}^{G}\sum_{j=1}^{G}p(i,j)$	2D, 3D
Second Order Difference Moment	$SDM=\sum_i\sum_j (i-j)^2 p(i,j)$	2D, 3D
First Order Difference Moment	$FDM=\sum_i\sum_j (i-j)p(i,j)$	2D, 3D
Second Order Inverse Difference Moment	$SIDM=\sum_i\sum_j \frac{1}{1+(i-j)^2}p(i,j)$	2D, 3D
Second Order Diagonal Moment	$SDIAGM=\sqrt{0.5*(j-i)p(i,j)}$	2D, 3D
Entropy	$ENTR=-\sum_i\sum_j p(i,j)*log(p(i,j))$	2D, 3D
Correlation	$CORR=\sum_i\sum_j \frac{(i*j)*P(i,j)-(u_x*u_y)}{\sigma_x*\sigma_y}$	2D, 3D
Peak Transition Probability	$PTP=max(p(i,j))$	2D
Uniformity	$UNI=\sum_i\sum_j p(i,j)^2$	2D, 3D
Gray Level Co-occurrence Matrix(GLRLM)		
Short Run Emphasis	$SRE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{j^2}$	3D
Long Run Emphasis	$LRE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)*j^2$	3D
Low Gray Level Run Emphasis	$LGRE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{i^2}$	3D
High Gray Level Run Emphasis	$HGRE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)*i^2$	3D
Short Run Low Gray Level Emphasis	$SRLGE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)}{i^2\cdot j^2}$	3D
Short Run High Gray Level Emphasis	$SRHGE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{p(i,j)*i^2}{j^2}$	3D
Long Run Low Gray Level Emphasis	$LRLGE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}\frac{P(i,j)\cdot j^2}{i^2}$	3D
Long Run High Gray Level Emphasis	$LRHGE=\frac{1}{n_r}\sum_{i=1}^{M}\sum_{j=1}^{N}p(i,j)*i^2*j^2$	3D
Gray Level Non-uniformity	$GLNU=\frac{1}{n_r}\sum_{i=1}^{M}\left(\sum_{j=1}^{N}p(i,j)\right)^2$	3D
Run Length Non-uniformity	$RLNU=\frac{1}{n_r}\sum_{j=1}^{N}\left(\sum_{i=1}^{M}p(i,j)\right)^2$	3D
Run Percentage	$RPC=\frac{n_r}{p(i,j)\cdot j}$	3D

two dimensions, co-occurrence matrices for volume data also represent an n*n matrix in which n is the gray level. These matrices are defined using the specific displacement vector $disp = (dx, dy, dz)$ for each direction, where dx, dy, and dz are the number of voxels that move along the x-, y-, and z-axis, respectively. With respect to each pixel, pixels in 26 directions can be examined, but only 13 directions were considered to avoid redundancy. From the calculated matrices, we extracted nine 3D texture features.

The basic concept used for expanding 3D GLRLM is very similar to that for 3D GLCM. Each component $p[i, j]$ of matrix p indicates the number of runs that have a gray-level value of i, and the length of the runs j in specific orientations. The size of matrix p can as expressed by m*k, where m and k are the maximum gray level and the length of the maximum run, respectively. From the calculated matrices, we extracted eleven 3D GLRLM texture features.

To confirm the validity of the 3D texture features, we also needed comparable classifier models. For this purpose, we selected four feature groups and extracted 25 2D features from 2D section data sets[15-17]. Table 1 and 2 list the entire 2D and 3D features that used for our validation.

Mayall/Young chromatin features are extracted

Table 2. The list of Mayall/Young chromatin, densitometric, and morphometric features

Mayall/Young Chromatin texture		
Heterogeneity	$HETERO = \frac{N_B + N_W}{A}$	2D
Clumpness	$CLUMP = \frac{\sum meshdiff}{N_B + N_W}$	2D
Condensation	$CONDENS = \frac{\sum meshdiff}{N_B + N_G + N_W}$	2D
Densitometric		
Gray Level Average	$AVG = \frac{1}{N}\sum_i\sum_j p(i,j)$	2D
Gray Level Variance	$VAR = \frac{1}{N}\sum_i\sum_j (p(i,j)-\mu)^2$	2D
Gray Level Standard Deviation	$STD = \sqrt{VAR}$	2D
Kurtosis	$KUTO = \frac{1}{N}\sum_i\sum_j (p(i,j)-\mu)^4$	2D
Skewness	$SKEW = \frac{1}{N}\sum_i\sum_j (p(i,j)-\mu)^3$	2D
Cluster Shade	$CS = (i+j-ui-uj)^3 p(i,j)$	2D
Cluster Prominence	$CP = (i+j-ui-uj)^4 p(i,j)$	2D
Coeff. of Variance	$CV = \frac{STD}{\mu} * 100$	2D
Morphomteric		
Area	$AREA = \sum$ (Pixels of Object)	2D
Perimeter	The length of the outside boun dary of the selection	2D
Feret's Diameter	The longest distance between two points along the selection boundary	2D
Circularity	$CIR = 4\pi(Area/Perimeter^2)$	2D

(A : Sum Area, N_B : Number of Black Pixel, N_W : Number of White Pixel, N_G : Number of Gray Pixels, $\sum meshdiff$ Number of differing pixels between white and black pixels in a check-board meshed tessellation of given width

from the 3-level image considering the global shape of the object. Three values regarding the extent of the condensation of chromatin used in the study include heterogeneity, clump, and condensation. B and W indicate the number of the white and black pixels, respectively, in each nucleus area. M, the size of mesh, was fixed at 8 pixel in this study. N_B, N_G, and N_W indicate the number of pixels that have a value of black, gray, and white, respectively, in each mesh area.

Heterogeneity of parameters is not sensitive to the distribution of condensed and non-condensed area. However, it emphasizes the difference between condensed and non-condensed chromatin area. For example, the parameter value becomes zero if the nucleus has a single homogeneous value at gray level, while the value becomes one in a completely heterogeneous gray level. Moreover, among the parameters regarding the granularity, such as the clump and condensation, the former is sensitive to the absolute size of the particles in nucleus. The low value indicates small size of the particle, and the particle smaller than the size of the mesh has a value close to zero. The latter indicates the ratio of the large particles to the total nucleus area.

Densitometric features are commonly used in image analysis and include various features, most of which are computed using statistical methods. In this study, a total of 8 densitometric features were used.

Morphometric features such as the size of the object, shape, and locations can be estimated using the geometrical features. Similar to the densitometric features, the features in this group have been commonly used as significant features for long. Since the area containing the normal or abnormal cells shows particular geometrical forms in histopathological images, such morphological features can be more useful than the statistical features in that the latter is more sensitive to the variance of the data.

2.4 Statistical Analysis

To verify the effectiveness of the 3D texture features, we performed a quantitative analysis involving the comparative study of several grading classifiers. Our test used 419, 430, 785, and 789 cell nuclei objects for grades 1 to 4, respectively. The 20 3D texture features were extracted from each of the 2,423 cell nuclei. First, we reduced the dimensionality of the features through a principal component analysis (PCA)[18-19]. Based on the result of this analysis, a total of five principal components were selected. Then, we used these five principal components as new features with the linear combination of each eigenvector and the original extracted feature values.

As the next step, we applied a discriminant analysis. We used a pre-performed training process to improve the correctness of the classification by using 100 training data sets that had been selected randomly from each grade. The rest of the nuclei data were used for the test. The correct classification rates using the 400 training data sets were 72, 87, 54, and 94% for grades 1 to 4, respectively. Finally, we created six classifiers (named A-F) and tested the correctness of grading. Table III shows the detailed specifications of each classifier.

Table 3. Specifications of the six classifiers

Classifier	Specifications
A	25 2D texture features without any statistical process
B	25 2D texture features with PCA
C	20 3D texture features with PCA
D	20 3D texture features without any statistical process
E	20 3D texture features and 3 3D morphological features without any statistical process
F	The classifier E with PCA

Classifiers E and F combined the 3D texture features with 3D morphological features (Volume, Surface Area, and Spherical shape Factor). As noted previously, the morphological characteristics of cell nuclei are still considered important factors for clinical grading. Moreover, in order to examine the discriminant power in classification of Grade 2 and Grade 3, which are known to be clinically difficult, the value of area under receiver operating curve (AUC) was investigated[20-21].

Finally, as a comparative study, we performed stepwise features selection to select another optimized features. Using the selected these features, we also inspect misclassification rate for each grade by statistical analysis in the same way[22].

Consequently, we expected to reduce the error rate of grading through these combinations. All statistical evaluations were made using the SAS program package (SAS Institute, Cary NC, USA).

3. RESULTS

A system with an Intel® Pentium®4 3.0 GHz processor and the Nvidia GeForce™ 6800XT graphics card was used for software implementation and the performance test. The software tool used for measuring 23 of the 3D features was newly implemented using Microsoft Visual C++® 6.0 (Microsoft, Redmond, WA, USA), OpenGL® library, and GLSL (OpenGL® Shading Language). The 20 2D features were measured using IMAN, an image analyzer developed at the Medical Image Technology Laboratory (MITL) of Inje University.

3.1 Volume visualization based on 3D texture mapping

Figure 2 shows an example of volume rendering. We also applied pseudo-color mapping to the volume object to improve the visibility of the cell nuclei. Figure 3 provides an example showing our volume visualization of cell nuclei for each grade. For grade 1, we can easily see the small size and

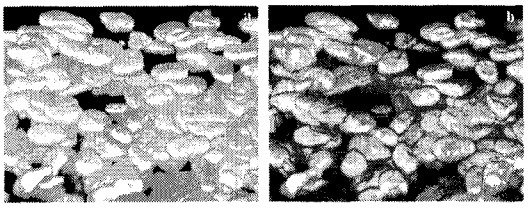


Fig. 2. An example of a rendered sample: (a) original gray-level and (b) pseudo-color mapping

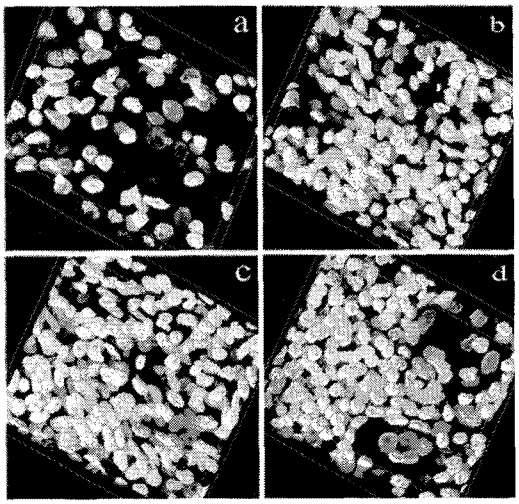


Fig. 3. Three-dimensional visualization of renal cell carcinoma cell nuclei: grades (a) I, (b) II, (c) III, and (d) IV

variation of cell shape. With progress from grade 2 to 3, the enlargement of cell nuclei is recognizable. Grade 4 shows a more pronounced impression of disorder, often with extreme variation. However, many clinical cases exist in which it is difficult to determine a grade based on a visual inspection of geometrical or morphological factors, such as the size of the cell nuclei. Therefore, the need for more objective, quantitative information is growing.

3.2 Classification

3.2.1 Classifier A and B

After extracting 25 features defined earlier, a principal components analysis is conducted to diminish the dimension of the features. Table 4 below

Table 4. The explanatory adequacy of 6 principal components

Eigenvalues of Correlation Matrix				
	Eigenvalue	Difference	Propotion	Cumulative
1	7.85685252	0.3347980	0.3022	0.3022
2	7.5220537	4.99617295	0.2893	0.5915
3	2.52588076	0.3578305	0.0971	0.6886
4	2.16859771	0.12961255	0.834	0.7721
5	1.03898515	0.09197585	0.0400	0.8120
6	0.94700930		0.0364	0.8484

presents an analysis result of the significant features from each principal component. Among 6 principal components chosen as default values, five principal components which explain about 81.20% were selected and redefined as new features. Accordingly, the eigenvectors were computed for 25 features and five new characteristic equations were defined by a linear combination of these eigenvectors and each of the features. After the principal components analysis, the validity of classification was checked by a discriminant analysis.

For a validity check, the result was compared with the one from using all of the 25 extracted 2D features without principal components analysis procedure. Two classifiers used in 2D analysis was named A and B were used for comparison between classifiers in the subsequent steps. Table 5 and Table 6 show the misclassification rate for each grade in the data sets. Comparing the result from using the five optimized features and the one using all of the 25 2D features, the difference in the accuracy of classification is clear.

Table 5. Misclassification rate of classifier A

Grade	1	2	3	4	Total
Rate	0.6923	0.1875	0.7143	0.5000	0.5235
Priors	0.250	0.250	0.250	0.250	

Table 6. Misclassification rate of classifier B

Grade	1	2	3	4	Total
Rate	0.2308	0.1250	0.5714	0.0714	0.2497
Priors	0.250	0.250	0.250	0.250	

3.2.2 Classifier C

Similar to the procedure for the classifier A and B, a principal components analysis was conducted to reduce the dimension of the features. Five selected principal components were then used as an optimized input value through a linear combination of the characteristic vector and the original features. After selecting the principal components and extracting the features, a discriminant analysis was conducted on the test data to create a classifier using the 3D feature. The classifier was created using the five new features that were diminished from 20 features and the function of the created classifier was evaluated. First, the classification rate for learning result using the 400 learning data was 72.00%, 87.00%, 54.00%, and 94.00% for each grade, with an average accuracy 76.75%. Applying the 2,023 test data to the created classifier after the learning completion, the classification accuracy for each grade turned out to be 77.13%, 71.61%, 63.94%, and 95.94%, with an average accuracy rate 78.30%. Table 7 below presents the number of data classified in each grade and the classification rate. For example, among 319 data that were previously categorized into Grade 1, only 72.73% of them, that

Table 7. Classification result for classifier C

Subj.G / CompG	1	2	3	4	Total
1	232 (72.73)	5 (1.57)	61 (19.12)	21 (6.58)	319 (100.00)
2	19 (5.76)	266 (80.61)	1 (0.30)	44 (13.33)	330 (100.00)
3	181 (26.42)	10 (1.46)	438 (63.94)	56 (8.18)	685 (100.00)
4	10 (1.45)	1 (0.15)	17 (2.47)	661 (95.94)	689 (100.00)
Total	442 (21.85)	282 (13.94)	517 (25.56)	782 (38.66)	2,023 (100.00)
Priors	0.25	0.25	0.25	0.25	

is, only 232 are correctly classified into Grade 1 according to the results from applying the classifier. Besides, five data, approximately 1.57%, were misclassified into Grade 2.

3.2.3 Classifier D

Classifier D is created using the 20 3D features as an input value without principal components analysis. It uses the previously extracted features without additional procedure. The classification accuracy was the lowest among the six classifiers, with an accuracy rate for each grade with the learning data 61.00%, 59.00%, 57.00%, and 50.00%, making the average accuracy rate 56.75%. The accuracy rate with test data was 45.45%, 34.24%, 43.36%, and 33.67%. With an average accuracy as low as 39.18%, the classifier turned out to be problematic.

3.2.4 Classifier E

Classifier E was created by adding three features of volume, surface area, and spherical shape factor to the twenty 3D features used in classifier D, and hence has both texture and morphological properties. The validity of those three features was already checked in a previous study that examines new grading system by a comparison analysis between morphological 3D features and2D features. These additional features play an important role in determining the grade in the actual clinical diagnosis.

After extracting three additional features for each object, the 23 features were used as an input value without the principal components analysis, similar to the procedure of classifier D. The accuracy rate with the learning data was 57.00%, 58.00%, 56.00%, and 51.00%, with an average accuracy 55.50%. The accuracy rate with test data was 44.83%, 37.58%, 41.29%, and 35.27%, with an average accuracy 39.49%, which is only 0.31% higher than that of the classifier D. Since the validity of the three morphological features was already checked, the overall classification accuracy was expected to be significantly improved. Nevertheless, the overall rate was similar to that of the classifier D.

3.2.5 Classifier F

Finally, classifier F is a modified type of classifier D. Classifier F uses 23 features used in classifier D as an input value and selects optimal five features through the principal components analysis which are subsequently used in the actual classification. The classification accuracy for the learning data was 94.00%, 89.00%, 75.00%, and 82.00%, with an average accuracy 82.50%. The classification accuracy for the test data was 91.22%, 86.45%, 76.50%, and 75.18%, with an average accuracy 82.19%. This is the best result among the six classifiers. Compared to the classifier C, even though the increase in accuracy was small with approximately 4.00%, it showed a stable accuracy with an even increase in every grade. This result indicates that the morphological analysis is an important factor in determining the classification. Combining the 3D texture used in this study with the morphological elements will provide a strong possibility for developing a superior classification system based on the 3D feature. Table 8 summarizes the results of applying classifier F to the test data.

Table 8. Classification result for classifier F

Subj.G / CompG	1	2	3	4	Total
1	291 (91.22)	0 (0.00)	27 (8.46)	1 (0.31)	319 (100.00)
2	2 (0.61)	282 (85.45)	5 (1.52)	41 (12.42)	330 (100.00)
3	161 (23.50)	0 (0.00)	524 (76.50)	0 (0.00)	685 (100.00)
4	3 (0.44)	164 (23.80)	4 (0.58)	518 (75.18)	689 (100.00)
Total	457 (22.59)	446 (22.05)	560 (27.68)	560 (27.68)	2,023 (100.00)
Priors	0.25	0.25	0.25	0.25	

3.2.6 Area under receiver operating curve analysis

Additionally, in order to find the discriminant power of the 3D texture feature for Grade 2 and Grade 3 images, 10 random sets were chosen whose area under receiver operating curve (AUC) was investigated. The nuclei in each dataset were categorized into Grade 2 and Grade 3 beforehand inside each set according to the properties of the nucleus through a visual inspection by the clinical diagnosis experts.

In general, the features that have discriminant power have AUC values between 0.5 and 1.0. Those with AUC>0.8 are considered to have a superior discriminant power. The table below presents the results of 10 dataset used in the test.

Some features showed a consistent discriminant power for each of the different dataset. Among the twenty features, entropy (ENTR) has the best discriminant power with the AUC larger than 0.5 for all data sets, followed by long run low gray level emphasis (LGRLE), second order inverse difference moment (SIDM), and second order diagonal moment (SDIAGM). Among the twenty features, about eight features turned out to have validity for Grade 2 and Grade 3, on average. In the table, the values that have high discriminant power with AUC higher than 0.8 is highlighted.

3.2.7 Classification after stepwise after selection

Finally, as an another approach, we performed

Table 9. AUC value table for 20 3D texture features

3D Feature \ #Set	1	2	3	4	5	6	7	8	9	10	AUC>0.5
ASM	0.426	0.396	0.432	0.601	0.543	0.433	0.598	0.353	0.245	0.455	3
CONT	0.455	0.513	0.667	0.479	0.494	0.367	0.456	0.312	0.433	0.470	2
SDM	0.399	0.432	0.325	0.435	0.462	0.411	0.501	0.467	0.333	0.251	1
FDM	0.300	0.531	0.398	0.487	0.349	0.444	0.547	0.673	0.349	0.513	3
SIDM	0.426	0.500	0.477	0.662	0.577	0.672	0.552	0.778	0.319	0.554	7
SDIAGM	0.761	**0.828**	**0.940**	0.782	**0.841**	0.452	**0.864**	0.345	0.303	0.742	7
ENTR	**0.835**	0.669	**0.851**	0.754	**0.812**	0.532	**0.901**	0.668	0.786	0.735	10
UNI	0.385	0.435	0.386	0.411	0.573	0.298	0.377	0.551	0.441	0.373	2
CORR	0.544	0.647	0.615	0.511	0.455	0.413	0.455	0.627	0.295	0.359	5
SRE	0.326	0.219	0.400	0.298	0.434	0.313	0.391	0.435	0.231	0.268	0
LRE	0.377	0.676	0.454	0.741	**0.812**	0.475	0.550	0.451	0.300	0.738	5
LGRE	0.783	**0.891**	0.642	0.677	**0.903**	0.412	0.523	0.742	0.512	**0.834**	9
HGRE	0.404	0.448	0.344	0.333	0.424	0.383	0.556.	0.341	0.272	0.452	1
SRHGE	0.686	0.549	0.556	0.431	0.613	0.414	0.647	0.406	0.319	0.499	5
SRLGE	0.587	0.453	0.666	0.564	0.448	0.452	0.454	0.563	0.400	0.342	4
LRHGE	0.448	0.643	0.418	0.325	0.431	0.339	0.462	0.489	0.345	0.377	1
LRLGE	0.485	0.521	0.555	0.411	0.385	0.546	0.477	0.544	0.237	0.416	4
GLNU	0451	0.434	0.319	0.676	0.364	0.370	0.358	0.338	0.398	0.327	1
RLNU	0.439	0.374	0.402	0.323	0.315	0.331	0.301	0.674	0.474	0.345	1
RP	0.713	**0.832**	0.695	0.672	0.784	0.398	0.762	0.590	0.349	0.460	4
#AUC>0.5	7	12	9	10	9	3	11	10	2	6	total
#Number of grade 2 nuclei	24	23	18	8	16	17	22	15	13	11	167
#Number of grade 3 nucle	15	9	7	24	18	21	18	9	14	23	158
#Total number of nuclei	39	32	25	32	34	38	40	24	27	34	325

Table 10. Stepwise selection result

step	Entered	F-Val.	Pr>F	Wilks' lambda	Pr<Wilks' lambda
1	LGRE	924.564	<.0001	0.29587	<.0001
2	ENTR	36.237	<.0001	0.10875	<.0001
3	SRHGE	4.563	<.0001	0.09231	<.0001
4	UNI	10.471	<.0001	0.07409	<.0001

stepwise features selection process with 20 3D texture features. From this process, we obtained 4 optimized features. Table 10 represent the final result of stepwise selection. Using these 4 features we performed a statistical classification in the same way and classification result was compared with others. The classification accuracy for the test data was 89.34%, 84.28%, 79.49%, and 84.16%, with an average accuracy 84.32%.

4. DISCUSSION

Figure 4 presents the classification accuracy of the six classifiers for each grade. When the optimized features were used from the PCA results, our classifier based on the 3D features generally showed a higher rate. In contrast, when the original feature values were used for classification, the rates were lower.

Nevertheless, using the optimized principal components as new features, the results showed relatively stable rates for each grade. When we compared the average rate between classifiers B and C, classifier C (78.30%) gave better results than B (75.13%). Comparing these two classifiers, except for grade 2, our new classifier C gave higher, more stable results overall. A more reliable result was obtained with the classifier F, which combined textural and morphological features. This gave the best result (82.19%), verifying the importance of morphological characteristics in clinical diagnosis. As for classifier F, its importance was confirmed in a previous study that tried the combination between morphological features, which is

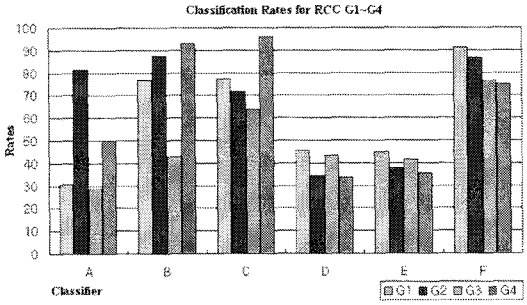


Fig. 4. Comparison of classification performance using six different classifiers

considered as the most important factor in actual clinical experiences. As was predicted, the accuracy was the highest with the rate slightly above 80.00%. This result opens a possibility for an alternative method that can improve the accuracy in classification by combining the optimized 3D texture features and 3D morphological features.

In the case of Grade 2 and Grade 3, which are known to be most difficult to classify in actual clinical experiences, the accuracy fluctuated from one classifier to another. When optimized principal components were used as feature, the classification rate turned out to be stable in each grade. Meanwhile, the misclassification rate of Grade 1 and Grade 4, which is relatively easy due to apparent characteristics, was turned out to be higher than that of Grade 2 and Grade 3. This probably has played a role in decreasing the overall classification rate. Following the cause analysis result, the problem in the composition of the learning data and optimization of the data are assumed to be the reason lying behind the low rate. It is surmised that the learning was not sufficient due to a relatively small size of extracted learning data for each group compared to the other studies. Considering the wide range of the actual extracted features according to the properties of the features, it would be beneficial to use some stereotype data.

When AUC was marginally conducted on a very small portion of the Grade 2 and Grade 3 data, a few features revealed consistently high discrim-

inant power.

Nevertheless, only half of the features turned out to have modest discriminant power with their overall AUC above 0.5 in each dataset. An improvement of accuracy for Grade 2 and Grade 3 is critical in enhancing the overall accuracy of the classification, and therefore a number of supplemental studies should be followed.

From the stepwise selection approach, we also could obtain more improved result. Compared with the classifier F, classification accuracy for grad 1 and 2 were lower, however the deviation for each grade were reduced and we could obtained most stable result. Figure 5 represents comparative performance result between the classifier F and the classifier using 4 stepwise selected features.

The 3D GLCM and GLRLM that were used in this study to extract the 3D feature are the extensions of the GLCM and GLRLM, which are traditional methods that are most commonly used in 2D texture analysis. Recently, a variety of methodologies regarding the extraction of 2D feature have been suggested. Hence, it is also necessary to consider the extension of these methods to three-dimension. Sufficient researches and data collection are also required to maintain the safety and robustness of the classifier. In addition to the statistical approach such as discriminant analysis used in this study, new research approach with an application of various recent classification models such as neural network, Fuzzy, SVM, etc. can also be considered in the future studies in order to improve the accuracy of classification, which is the most important factor in considering the classifiers.

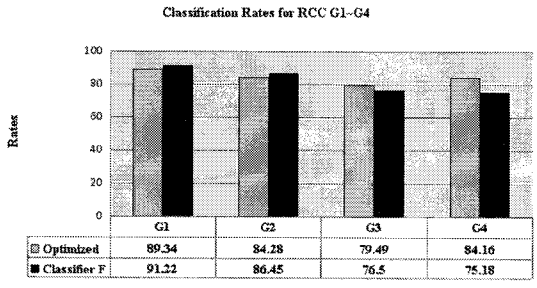


Fig. 5. Comparison of classification performance between the classifier F and the classifier using 4 stepwise selected features.

5. CONCLUSIONS

In this study, the image of cell nuclei for RCC obtained by CLSM was reconstructed into volume data and was subsequently visualized through 3D surface rendering. After the newly defined 3D texture features were extracted, new classifiers were created whose accuracy was compared with that of the classifier which uses the previous 2D feature. The validity was checked using 3D features in diverse forms.

According to the statistical treatment and analysis results, the accuracy has improved compared to the previous classification rate based on the 2D feature. Moreover, through a continuous study in the future, a further improvement in accuracy is expected by combining the 3D texture features with other 3D features such as the carcass features and refining them. This study provides a foundation for creating a new classification system with superior function in future studies.

The new classification method, which uses the morphological analysis through the 3D visualization and quantified 3D texture information, is expected to overcome the limitations of the previous system based on the 2D image analysis. It can be used as a research tool that can contribute to reducing misdiagnosis of RCC (Renal Cell Carcinoma).

REFERENCES

[1] R.F. Walker and P.T. Jackway, "Statistical Geometric Features Extension for Cytolo gical Texture Analysis," *IEEE Proceeding of the ICPR'96*, pp. 790-794, 1996.

[2] C. François et al., "The Chromatin Pattern of Cell Nuclei is of Prognostic Value for in Renal

Cell Carcinoma," *Analytical Cellular Pathology*, Vol.16, pp. 161-75, 1998.

[3] C. Roussile, S. Paillasson, and M.R. Nicoud, "Chromatin Texture Analysis in Living Cell," *The Histochemical Journal*, Vol.31, pp. 63-70, 1999.

[4] Q. Ji, J. Engel, and E. Craine, "Texture Analysis for Classification of Cervix Lesions," *IEEE Transactions on Medical Imaging*, Vol.19, pp. 1144-1149, 2000.

[5] G. Wouwer et al., "Wavelet as Chromatin Texture Descriptors for the Automated Identification of Neoplastic Nuclei," *Journal of Microscopy*, Vol.197, pp. 25-35, 2000.

[6] T. Irinopoulou et al., "Three-dimensional DNA Image Cytometry by Confocal Scanning Laser Microscopy in Thick Tissue Blocks of Prostatic Lesions," *Cytometry*, Vol.27, pp. 99-105, 1997.

[7] K. Jafari-Khouzani et al., "Comparison of 2D and 3D Wavelet Features for TLE Lateralization," *Proceeding of the SPIE Medical Imaging 2004 : Physiology, Function and Structure from Medical Images*, Vol.5369, pp. 593-601, 2004.

[8] A. Madabhushi et al., "A Novel Stochastic Combination of 3D Texture Features for Automated Segmentation of Prostatic Adenocarcinoma from High Resolution MRI," *Medical Image Computing and Computer-assisted Intervention*, Vol.2878, pp. 581-591, 2003.

[9] A.S. Kurani et al., "Co-occurrence Matrices for Volumetric Data," *7th IASTED Int'l Conference on Computer Graphics and Imaging*, Kauai, Hawaii, USA, in August 16-18, 2004.

[10] D.H. Xu et al., "Run-length Encoding for Volumetric Texture," *4th IASTED Int'l Conference on Visualization, Imaging and Image Processing*, Marbella, Spain, September 6-8, 2004.

[11] H.J. Choi et al., "Three-dimensional Visualization and Quantitative Analysis of Cervical Cell Nuclei with Confocal Laser Scanning Microscopy," *Analytical and Quantitative Cytology and Histology*, Vol.27, No.3, pp. 174-180, 2005.

[12] T.Y. Kim et al., "3D Quantitative Analysis of Cell Nuclei based on Digital Image Cytometry," *Journal of Korea Multimedia Society*, Vol.10, No.7, pp. 846-855, 2007.

[13] R.J. Rost, *OpenGL® Shading Language*, Addison Wesley, Boston, 2004.

[14] M. Meißner, U. Hoffmann, and W. Straßer, "Enabling Classification and Shading for 3D Texture Mapping based Volume Rendering Using OpenGL® and Extentions," *Proceeding of the 10th IEEE Visualization 1999 Conference (VIS'99)*, pp. 207-214, 1999.

[15] K. Rodenacker and E. Bengtsson, "A Feature Set for Cytometry on Digitized Microscopic Images," *Analytical Cellular Pathology*, Vol.25, pp. 1-36, 2003.

[16] C. Wahlby, "Algorithms for Applied Digital Image Cytometry," Ph.D. thesis, Uppsala University, Sweden. 2003.

[17] I.T. Young, P.T. Verbeek, and B.H. Mayall, "Characterization of Chromatin Distribution in Cell Nuclei," *Cytometry*, Vol.7, pp. 467-474, 1986.

[18] R.A. Johnson, *Applied Multivariate Statistical Analysis 5th Ed.*, Prentice-Hall, Upper Saddle River, NJ, 2002.

[19] R. Johnsonbaugh and S. Jost, *Pattern Recognition and Image Analysis*, Prentice-Hall, Upper Saddle River, NJ, 1996.

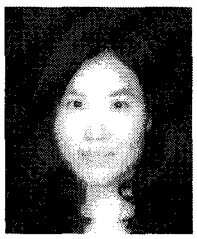
[20] T. Fawcett, "An Introduction to ROC Analysis," *Pattern Recognition Letters*, Vol.27, pp. 861-874, 2006.

[21] V. Bewick, L. Cheek, and J. Ball, "Statistics Review 13 : Receiver Operating Characteristics Curves," *Critical Care*, Vol.8, No.6, pp. 508-512, 2004.

[22] H. Schulerud et al., "A Review of Caveats in Statistical Nuclear Image Analysis," *Analytical Cellular Pathology*, Vol.16, No.2, pp. 63-82, 1998.

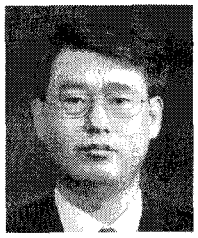
Tae-Yun, Kim

He received the M.S. degree in computer science at the department of computer science of Inje University in 2006. He is currently a Ph.D. candidate at Inje university. His interesting research fields are in pathological image analysis, medical image processing, and computer graphics.



Hyun-Ju, Choi

She is currently working at the BK21 Biomedical Science Education Center, School of Medicine, Pusan National University, as a Research Assistant Professor. She received her Ph.D. in Computer Science from Inje University, South Korea in 2005. Her research interest is computerized two- and three-dimensional image analysis using cells and tissues in vivo as well as in vitro.



Heung-Kook, Choi

He has gone the undergraduate studying and graduate studying in computer science and engineering at the department of electrical engineering of Linköeping University, Sweden (1984-1990) and Ph.D. studying in computerized image analysis at the Center for Image Analysis of Uppsala University, Sweden(1990-1996). Now he is president of Industry and Academic Cooperation Foundation at Inje University and vice president of Korea Multimedia Society. His interesting research fields are in computer graphics, virtual reality, medical image processing, and medical image analysis.