Multiwavelet grading of prostate pathological images

Hamid Soltanian-Zadeh[*a,b], Kourosh Jafari-Khouzani[**c]
[a]Radiology Image Analysis Lab, Henry Ford Health System, Detroit
[b]Dept. of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[c]Dept. of Computer Science, Wayne State Univ., Detroit

ABSTRACT

We have developed image analysis methods to automatically grade pathological images of prostate. The proposed method generates Gleason grades to images, where each image is assigned a grade between 1 and 5. This is done using features extracted from multiwavelet transformations. We extract energy and entropy features from submatrices obtained in the decomposition. Next, we apply a *k*-NN classifier to grade the image. To find optimal multiwavelet basis, preprocessing, and classifier, we use features extracted by different multiwavelets with either critically sampled preprocessing or repeated row preprocessing and different *k*-NN classifiers and compare their performances, evaluated by total misclassification rate (TMR). To evaluate sensitivity to noise, we add white Gaussian noise to images and compare the results (TMR's). We applied proposed methods to 100 images. We evaluated the first and second levels of decomposition using Geronimo, Hardin, and Massopust (GHM), Chui and Lian (CL), and Shen (SA4) multiwavelets. We also evaluated *k*-NN classifier for *k*=1,2,3,4,5. Experimental results illustrate that first level of decomposition is quite noisy. They also show that critically sampled preprocessing outperforms repeated row preprocessing and has less sensitivity to noise. Finally, comparison studies indicate that SA4 multiwavelet and *k*-NN classifier (k=1) generates optimal results (with smallest TMR of 3%).

Keywords: Gleason grading system, multiwavelet transform, *k*-NN classifier, simulated annealing, image analysis, classification

1. INTRODUCTION

1.1 Background and motivation

Cancer is the second killer of American people, and only cardiovascular diseases take a higher toll.[1] Histological grading is a very important task in the framework of prostate cancer prognosis, since it is used for treatment planning. If infection of cancer disease was not rejected by non-invasive diagnostic techniques like MRI, CT scan, and ultrasound, then a biopsy specimen of the tissue is tested. For prostate, the tissue is usually stained by H&E (Hematoxyline and Eosine) technique. Then the histological grading is done by viewing the microscopic image of the tissue. This task is done by pathologists. Manual grading is very subjective due to inter- and intra-observer variations. So an automatic and repeatable technique is needed for grading. Gleason grading system is the most common method for histological grading of prostate.[2] The goal of this paper is to automate the Gleason grading.

For data classification, the decision is made based on a set of features. Since most pattern recognition tasks are first done by humans and are automated later, the most appropriate source of features has been those used by the experts to classify the objects. Automating the classification of objects using the same features as those used by experts can be a difficult task, but fortunately the features used by machines need not be precisely those used by humans. Sometimes features that would be impossible or difficult for humans to estimate are useful in automated systems.[3] In this research, we used energy and entropy features calculated from multiwavelet coefficients of the image. Then a *k*-NN classifier was used to classify each image to appropriate grade. The leaving-one-out technique was used for error rate estimation.

* hamids@rad.hfh.edu; phone 1 313 874-4482; fax 1 313 874-4494; http://Radiologyresearch.com; Radiology Image Analysis Lab, One Ford Place, 2F, Detroit, Henry Ford Health System, Detroit, MI48202, USA; ** kjafari1@yahoo.com; phone 1 313 577-5070; fax 1 313 577-6868; www.cs.wayne.edu; Dept. of Computer Science, 431 State Hall, Wayne State Univ., Detroit, MI48202, USA

1.2 Gleason grading system

There is a great need for methods to quantify the probable clinical aggressiveness of a given neoplasm, and further to express its apparent extent and spread in patients.[1] Histological grading is one of these methods. The grading of a cancer attempts to establish some estimate of its aggressiveness or level of malignancy. In Gleason grading system, the prostate cancer may be classified as grade 1, 2, 3, 4 or 5 with increasing or lack of glands differentiation as explained below.

Gleason has provided a conceptual diagram in Fig. 1 to show the continuum of deteriorating cancer cell architecture, and the four dividing lines along this continuum that he discovered are able to identify patients with significantly different prognosis. The Gleason system is based exclusively on the architectural pattern of the glands of the prostate tumor. It evaluates how effectively the cells of any particular cancer are able to structure themselves into glands resembling those of the normal prostate.[2] The ability of a tumor to mimic normal gland architecture is called its *differentiation*, and experience has shown that a tumor whose structure is nearly normal (well differentiated) will probably have a biological behavior relatively close to normal (that is not very aggressively malignant). Gleason grading from very well differentiated (grade 1) to very poorly differentiated (grade 5) is usually done by viewing the low magnification microscopic image of the prostate tissue.

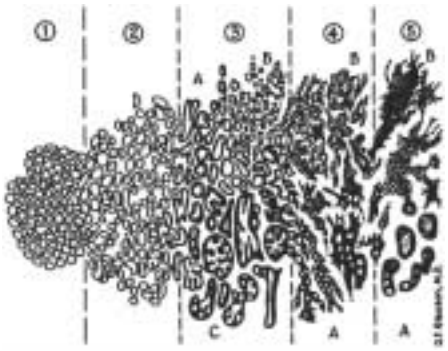
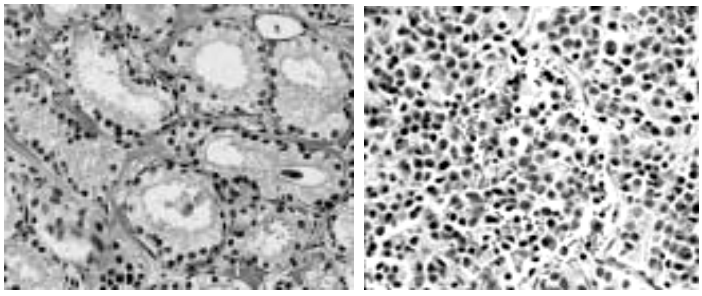


Figure 1. Gleason grading diagram.

If there exits two patterns in the specimen, a combined score is calculated which is the sum of two grades. So combined score varies from 2 to 10. Fig. 2 shows two tissue samples of grades 2 and 5. For grade 2, the glands are well-differentiated with respect to grade 5. Fig. 2(b) shows only a sea of black nuclei with no pattern.

The grade of a prostate cancer specimen is very valuable to doctors in understanding how a particular case of prostate cancer can be treated. An accurate Gleason score can help one decide which treatment may be most beneficial. In general, the time for which a patient is likely to survive following diagnosis of prostate cancer is related to the Gleason score. The lower the Gleason score, the better the patient is likely to do. Patients with score of 2 to 4 almost never develop aggressive disease, whereas most patients with a score of 8 to 10 die of prostatic carcinoma.[2]



(a) (b)

Figure 2. Two samples of prostate tissue. (a) Grade 2. (b) Grade 5.

1.3 Previous work

Analysis of pathological images has been an area of interest during the last few years.[4-13] The aim of these researches has been distinguishing the normal and abnormal tissues. Stotzka *et al.*[4] proposed a method to distinguish the moderately and poorly differentiated lesions of prostate tissues. The decision is based on a number of features based on shape and texture of the image. In reference[5] a technique using pyramid node linking to segment and classify the given cell images is described. The proposed method was used for microscope slides of cultured rat liver cells, to classify these cells into one of three possible classes. The decision is based on previous knowledge of gray levels in these groups. A method for automatic grading of breast cancer based on Bloom and Rechardson grading system has been proposed in reference[6]. Features based on fuzzy co-occurrence matrix are calculated and then the decision is made using an artificial neural network. A nonlinear technique is proposed in reference[7] to segment and extract features from the area of each individual cell in biopsy images of breast. Then a fuzzy classifier is used which determine the probability of the biopsy to belong to a high or low cancer level.

Schnorrenberg *et al.*[8] developed a method to detect tissue cell nuclei in histological sections of breast with immuno-cytochemistry staining. The detection system uses a receptive field filter to enhance negatively and positively stained cell nuclei and a squashing function to label each pixel value as belonging to the background or a nucleus. Some statistical features of color values of cell pixels are calculated, and then a neural network is used to classify each cell to one of five classes.

An automatic system was presented in reference[9] to analyze a cell nucleus in a given biopsy of mammary tissue, which is cancerous. Images are enhanced and segmented using morphological transformations. An ultimate erosion is used in two steps to separate cell nuclei in contact. It is based on a combination of symmetrical ultimate erosion with directional ultimate erosion. Wouwer[10] has proposed a method for classification of pathological images of breast. Features based on wavelet transform are extracted from each segmented cell and the cell is classified to one of four grades. The grading is based on distribution of chromatin in the cell.

Hallinan proposed a method for detection of malignancy in cervical cells[11]. The cytological image is first segmented to cells. A number of features are defined for each cell. Then an artificial neural network is used for classification. A similar study was done in reference[12] to determine malignant mesothelioma. A number of shape features are calculated for each nucleus and a *k*-NN classifier is used for classification. An automatic algorithm for the categorization of normal and cancerous colon mucosa was reported in reference[13] where a number of features were derived using the co-occurrence matrix and a parametric linear-discriminate function was used to determine the classification rule.

A major difference between our work and most previous techniques is that they use the shape information of individual cells or glands and/or its texture information, but we use features of the entire image. Another difference is that we use multiwavelets which have not been used in previous work.

2. FEATURE EXTRACTION

2.1 Multiwavelet transform

While in scalar wavelet transform there is only one scaling function, in multiwavelet transform we can have more than one scaling function. Multiwavelets have some advantages compared to scalar ones. For example, features such as short support, orthogonality, symmetry and vanishing moments are known to be important in signal and image processing. A scalar wavelet cannot possess all of these properties at the same time. On the other hand, a multiwavelet system can simultaneously provide perfect reconstruction while preserving length (orthogonality), good performance at the boundaries (via linear-phase symmetry), and a high order of approximation (via vanishing moments). This suggests that multiwavelets may perform better in various applications.[14]

In multiwavelet analysis, the multiscaling function $\boldsymbol{\Phi}(t) = [\phi_1(t),...,\phi_r(t)]^{\mathrm{T}}$ satisfies a two-scale equation:

$$\boldsymbol{\Phi}(t) = \sqrt{2}\sum_k H_k \boldsymbol{\Phi}(2t-k) \tag{1}$$

where H_k is an $r \times r$ matrix of lowpass filter coefficients and r is called multiplicity. Like scalar wavelet function, multiwavelet function $\boldsymbol{\Psi}(t) = [\psi_1(t),...,\psi_r(t)]^T$ must satisfy the two-scale wavelet equation:

$$\boldsymbol{\Psi}(t) = \sqrt{2}\sum_k G_k \boldsymbol{\Phi}(2t-k) \tag{2}$$

where G_k is an $r \times r$ matrix of highpass filter coefficients.

Corresponding to each multiwavelet system is a matrix-valued multirate filterbank, or multifilter shown in Fig. 3. The lowpass filter and highpass filter consist of coefficients corresponding to the dilation equation (1) and wavelet equation (2) and these coefficients are matrices, so during the convolution step they must multiply vectors (instead of scalars). This means that multifilter banks need input rows. Thus, some methods for vectorization of scalar input should be used. These are called preprocessing methods and different approaches to preprocessing have been developed.[15-17] In this research, we use repeated row and critically sampled approaches.

In repeated row approach the input signal is repeated to get an input vector.[14] So it introduces oversampling of the data by a factor of two. There is an alternative version of repeated row preprocessing in which the first row of input vector is the signal and the second row is the signal with a factor of α. This factor is chosen so that if the input signal is constant, then the output from the high-pass multifilter is zero.[18] We use this kind of repeated row preprocessing.

In critically sampled approach the input signal is preprocessed such that a critically sampled representation is maintained: If the data enters at rate R/2, preprocessing yields two streams at rate R/2 for input to the multifilter.[14] The symmetric extension of signal also is also used as described in reference[14] to preserve critically sampling nature of system in filtering the signals at their boundaries. This approach can be used for symmetric or antisymmetric filter banks. All the multiwavelets that we used in this research have symmetric or antisymmetric filter banks except cardinal balanced multiwavelets which do not need preprocessing.

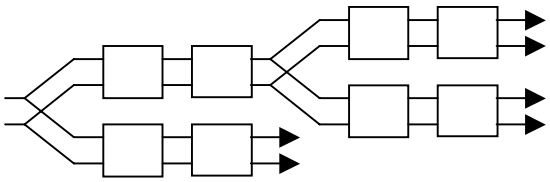


Figure 3. Multiwavelet filterbank, showing 2 levels of decomposition.

2.2 Multiwavelet transform of images

For calculating multiwavelet transform images, we can use tensor product method, i.e., performing the 1-D algorithm in each dimension separately.[14,17] Fig. 4 shows the submatrices resulted from 2-D multiwavelet decomposition. The result after first decomposition can be realized as the following matrix:

$$\begin{matrix} L_1L_1 & L_2L_1 & H_1L_1 & H_2L_1 \\ L_1L_2 & L_2L_2 & H_1L_2 & H_2L_2 \\ L_1H_1 & L_2H_1 & H_1H_1 & H_2H_1 \\ L_1H_2 & L_2H_2 & H_1H_2 & H_2H_2 \end{matrix}$$

in which each entry represents a subband, which corresponds to lowpass and/or highpass filters used in vertical and horizontal directions. For example, the subband labeled L_1H_2 corresponds to data obtained by applying the highpass

filter on the horizontal direction and taking its 2nd channel, then applying lowpass filter on the vertical direction and taking its first channel (Fig. 3). The next level of decomposition will decompose the following "low-low pass" submatrix, in a similar manner:

$$\begin{matrix} L_1L_1 & L_2L_1 \\ L_1L_2 & L_2L_2 \end{matrix}$$

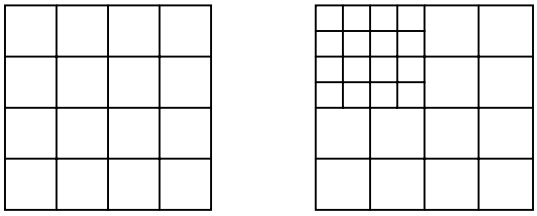
This is shown in Fig. 4(b). The number of submatrices will be equal to 4+12*l* where *l* is the number of levels of decomposition. The energy and entropy of the multiwavelet coefficients are calculated as features for image classification. As indicated in Fig. 4, the result of decomposition is a number of submatrices. From each submatrix $[x_{ij}]$, the following features are calculated:

$$Energy = \frac{\sum_i \sum_j {x_{ij}}^2}{N \times N} \tag{3}$$

$$Entropy = \frac{-1}{\log N^2} \sum_i \sum_j \left[\frac{{x_{ij}}^2}{norm^2} \right] \log \left[\frac{{x_{ij}}^2}{norm^2} \right] \tag{4}$$

where $norm^2 = \sum_i \sum_j {x_{ij}}^2$ and *N* is the dimension of each submatrix and its use in the above equation permit features be independent of submatrix dimensions.

In this work we use 10 different multiwavelets: GHM, CL, SA4, BiGHM2, BiH52s, BiH32s, BiH54n, CardBal2, CardBal3 and CardBal4. A brief description of each multiwavelet is given in Appendix.



(a) (b)

Figure 4. Result of 2-D multiwavelet decomposition. (a) One level of decomposition. (b) Two levels of decompostion.

3. CLASSIFICATION

3.1 *k*-NN classifier

Having generated a feature vector for each image, we use a *k*-nearest neighbors (*k*-NN) classifier using Euclidean distance to classify it to appropriate grade. Because limited number of images for each grade was available, we use the leaving-one-out technique to evaluate accuracy of classification. Before classification, we normalize features. Recall that if one of the features has a very wide range of possible values compared to the other features, it will have a huge effect on the dissimilarity (distance), and the decisions will be based primarily upon this single feature. To overcome this, it is necessary to apply scale factors to the features before computing the distances.[3] In this research, we normalized each feature to have mean of zero and standard deviation of one for the entire data set. Furthermore, because some features may be more important than others, we used weight for each normalized feature. To calculate the best weight vector for the feature vector, we minimized the error rate estimated by the leaving-one-out technique. This minimization was done using Simulated Annealing algorithm.[18] The *k*-NN algorithm classifies each image by assigning it the label most frequently represented among the *k* nearest samples; in other words a decision is made by examining the labels on the *k* nearest neighbors and taking a vote. If the label coincides with the Gleason grade of the sample, this is considered a correct classification

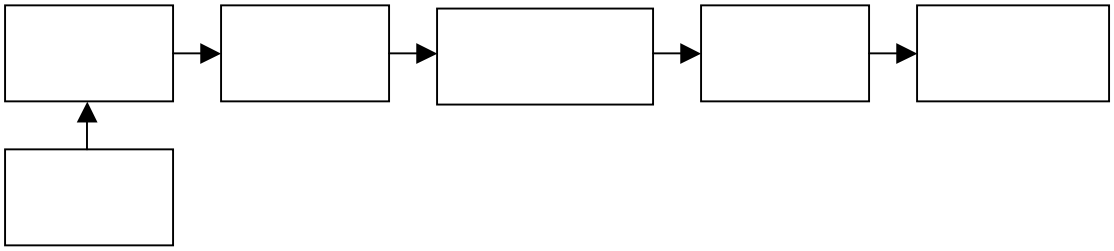


Figure 5. Stages of prostate image analysis: Image acquisition, generating black and white images, feature extraction, normalization, exerting weight vector, and classification.

3.2 Noise effect

To evaluate the noise effect we added Gaussian noise with signal to noise ratio SNR=5 to each image before classification in leaving-one-out technique. SNR is defined by the ratio of signal expectation to the noise standard deviation.

4. RESULTS

In our experiments, 100 prostate tissue sample images were processed by the proposed approach. These color images were of grades 2 to 5 and of magnification ×100 with different sizes.The specimens were stained using Hematoxyline and Eosine technique. Grade 1 was excluded because it is a very rare pattern and should be avoided. All of the images were captured in equal conditions of light. Our image set consisted of 21, 20, 32, 27 images of grades 2, 3, 4, 5, respectively. Because different portions of a specimen may have more than one grade, we captured our images in a manner so that each image has a single grade. The grading of these images was done by pathologists.

We first made each image black and white to simplify the calculations, since the color does not have important information. For multiwavelet features, each image was decomposed to submatrices as explained in Section 2.2. A set of features using relations (3) and (4) was calculated using submatrices, and then normalized as described in Section 3.1. Submatrices of the first and second levels of decomposition were tested separately using GHM, CL and SA4 multiwavelets. For each set of features, the *k*-NN classifier was tested for k=1, 3, 5 and 7, and the error was calculated using leaving-one-out technique. Tables 1-3 show the error rates before and after using weight vectors. In these tables, r.r. and c.s. show repeated row and critically sampled preprocessing, respectively.

Likewise because of similar results, we evaluated the effect of noise for only these three multiwavelets. The results of noise effect are given in Tables 1-3. These results are the average of error for 10 realization of Guassian noise. The results are rounded. We can see that the first level of decomposition is very sensitive to noise and should be avoided for feature extraction. This also helps for noise reduction. Furthermore for the first level of decomposition, critically sampled preprocessing leads to more errors compared to repeated row technique. This is because critically sampled preprocessing creates a compact form, so the coefficients resulted from signal in first level of decomposition are small compared to noise. As a result the SNR in this level is low, leading to a higher noise sensitivity.

As we see, for the second level of decomposition, critically sampled preprocessing has lower sensitivity to noise compared to repeated row preprocessing. This is also due to compact form that critically sampled technique can produce. This leads to higher energy and higher SNR at low resolutions and resulting in sensitivity to noise.

To compare different multiwavelets, we calculated the errors for the second level of decomposition using 10 different of multiwavelets with critically sampled preprocessing[1] and leaving-one-out technique to calculate the error rate. The results are graphed in Fig. 6. The results show that the multiwavelet basis determines the classification accuracy. As shown in Fig. 6, SA4 multiwavelet shows better results compared to other multiwavelets. Note that for high *k*'s the error

[1] As described in Appendix, cardinal multiwavelets (CardBal2, CardBal3, CardBal4) do not need preprocessing step.

grows rapidly. This is because of the number of data points is small compared to the number of classes. This causes that in the feature space there are not enough neighbors for an image with the same class.

5. CONCLUSIONS

We proposed a method for grading the pathological images of prostate. The color image is converted to black and white and then decomposed to multiwavelet submatrices. For each submatrix, the energy and entropy features are calculated and then normalized. Weight vectors are exerted and classification is done using *k*-NN classifier. The weight vectors are found using Simulated Annealing algorithm. This study demonstrates that energy and entropy features drived from multiwavelet transform can result in accurate classification and discrimination of various cancer grades in pathological images of prostate. It was shown that the multiwavelet basis affects the classification. The second level of decomposition has less sensitivity to noise. Because in this level the SNR is more than the first level. Furthermore for second level of decomposition, critically sampled preprocessing leads to less sensitivity compared to repeated row preprocessing. This is because critically sampled preprocessing creates better compactness.
One of the drawbacks of multiwavelets in feature extraction is the large number of produced features. Coarser resolutions may have important information, but with higher decomposition levels, the number of submatrices grows rapidly. In future studies, we are planning to reduce the dimention of the feature space.

Table 1. Error rates for first and second levels of decomposition using GHM multiwavelet.

		1st level				2nd level			
k		1	3	5	7	1	3	5	7
r.r.	Without Weight	20	28	31	40	16	22	28	25
	With Weight	13	16	22	28	14	17	20	20
	With Noise	26	28	27	36	33	31	30	36
c.s.	Without Weight	12	12	17	22	11	12	18	26
	With Weight	8	9	13	18	6	9	14	20
	With Noise	43	51	34	34	11	14	16	21

Table 2. Error rates for first and second levels of decomposition using CL multiwavelet.

		1st level				2nd level			
k		1	3	5	7	1	3	5	7
r.r.	Without Weight	17	27	32	35	10	18	17	24
	With Weight	13	18	20	27	8	11	13	16
	With Noise	30	29	30	37	20	18	18	26
c.s.	Without Weight	11	11	15	19	11	12	14	26
	With Weight	7	9	13	15	6	10	12	19
	With Noise	36	43	34	41	11	12	15	25

Table 3. Error rates for first and second levels of decomposition using SA4 multiwavelet.

		1st level				2nd level			
k		1	3	5	7	1	3	5	7
r.r.	Without Weight	17	23	26	27	16	22	29	29
	With Weight	11	15	19	18	15	17	19	24
	With Noise	25	34	36	34	53	44	55	57
c.s.	Without Weight	10	10	16	18	9	11	15	25
	With Weight	5	6	11	14	3	7	10	17
	With Noise	77	73	73	84	5	11	11	21

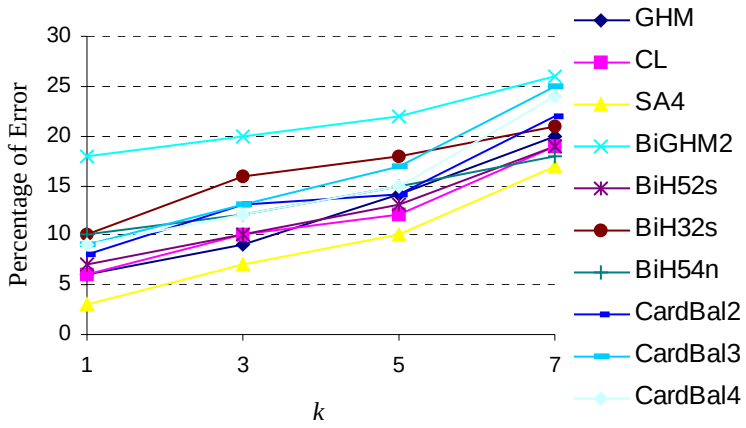


Figure 6. Comparison of error results, using different multiwavelet transforms. The features were extracted using second level of decomposition and critically sampled preprocessing.

6. APPENDIX

1) GHM: This multiwavelet was introduced by Geronimo, Hardin and Massopust. Both scaling functions are symmetric and multiwavelet functions are symmetric-antisymmetric.[19] It has approximation order of 2.
2) CL: This multiwavelet was introduced by Chui and Lian[20] and has approximation order of 2.
3) SA4: Shen *et al.*, showed how to create symmetric-antisymmetric orthonormal multiwavelets from orthonormal scalar wavelets.[21] Then they obtained the SA4 multiwavelet with length 2 from Daubechies scalar wavelets with length 4.
4) BiGHM2: This is a biorthogonal multiwavelet with length 2. Strela suggested a method to design biorthogonal multiwavelets with desired approximation order from ordinary multiwavelets.[22] BiGHM2 multiwavelet was made from GHM multiwavelet using this method.
5) BiH52s: This is a symmetric-antisymmetric biorthogonal multiwavelet with length 5 and approximation order of 2. It was proposed by Turcajova using Hermite splines.[23]
6) BiH32s: This multiwavelet is the dual of BiH52s.[23]
7) BiH54n: This is a biorthogonal multiwavelet with length of 5 and approximation order of 4. It can be obtained with the same proposed method in reference[22] but starting with Hermite multiwavelet.[23]
8) CardBal2: Cardinal multiwavelets were introduced to avoid the prefiltering step in multiwavelet computations.[24] Multiwavelet bases, for which the zero moment properties carry over to the discrete-time filter bank, are called balanced.[25] CardBal2 is a cardinal balanced multiwavelet introduced in reference[26] with length of 6 and approximation order of 2.
9) CardBal3: Cardinal balanced multiwavelet with length 8 and approximation order of 3 introduced in reference.[26]
10) CardBal4: Cardinal balanced multiwavelet with length 12 and approximation order of 4 introduced in reference.[26]

REFERENCES

1. Kumar, Cotran, Robbins, *Basic Pathology*, W.B. saunders company, 1997.
2. J. Rosai, L.V. Ackerman, *Ackerman's Surgical Pathology*, Mosby Inc., 1996.
3. E. Gose, R. Johnsonbaugh, and S. Jost, *Pattern Recognition and Image Analysis*, Prentice Hall, 1996.
4. Stotzka R., Männer R., Bartels P.H., and Thompson D. "A hybrid neural and statistical classifier system for histo-pathologic grading of prostate lesions," in *Analytical and Quantitative Cytology and Histology*, **17(3)**, pp. 204-218, 1995.
5. F. Arman, and J.A. Pearce, "Unsupervised classification of cell images using pyramid node linking," *IEEE Trans. Biomed. Eng.*, **37(6)**, pp. 647-650, June 1990.

6. H. D. Cheng, C. H. Chen, and R. I. Freimanis, "A neural network for breast cancer detection using fuzzy entropy approach," in *Proc. Int. Conf. Image Processing*, **3**, pp. 141-144, 1995.
7. S. M. Marroquin, C. Vos, E. Santamaria, X. Jove, and J.C. Socoro, "Non linear image analysis for fuzzy classification of breast cancer," in *Proc. Int. Conf. Image Processing*, **2**, pp. 943-946, 1996.
8. F. Schnorrenberg, C. S. Pattichis, K. C. Kyriakou, and C. N. Schizas, "Computer-aided detection of breast cancer nuclei," *IEEE Trans. Infor. Tech. in Biomed.*, **1(2)**, pp. 128-140, June 1997.
9. E. M. Marroquin, E. Santamaria, X. Jove, and J. C. Socoro, "Morphological analysis of mammary biopsy images," in *Electrotechnical Conf., 1996. MELECON '96., 8th Meditarranean*, **2**, pp. 1067-1070.
10. G. Van de Wouwer, "Wavelet for Multiscale Texture Analysis," *Ph.D. Thesis*, University of Antwerpen, Dept. Natuurkunde, 1998.
11. J. S. Hallinan, "Detection of malignancy associated changes in cervical cells using statistical and evolutionary computation techniques," *Ph.D. thesis*, *The University of Queensland, Australia*, 1999.
12. B. Weyn, G. V. Wouwer, Samir Knmar-Singh, A. Van Daele, Paul Scheunders, Eric Van Marck, and Willem Jacob, "Computer-assisted differential diagnosis of malignant mesothelioma based on syntactic structure analysis", Cytometry, **35**, pp. 23-29, 1999.
13. A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, and A. Murray, "Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa", *IEEE Trans. on Information Technology in Biomedicine*, **2(3)**, pp. 197-203, 1998.
14. V. Strela, P. Heller, G. Strang, P. Topiwala, and C. Heil, "The application of multiwavelet filter banks to signal and image processing," *IEEE Trans. Image Processing*, **8(4)**, pp. 548-563, 1999.
15. X.G. Xia, J.S. Geronimo, D.P. Hardin, and B.W. Suter, "Design of prefilters for discrete multiwavelet transforms," *IEEE Trans. Signal Processing*, **44**, pp. 25-35, 1996.
16. D.P. Hardin and D.W. Roach, "Multiwavelet prefilters I: Orthogonal prefilters preserving approximation order p<=2," *IEEE Trans. Circuits and Systems*, **45(8)**, pp. 1106-1112, Aug. 1998.
17. V. Strela and A. T. Walden, "Signal and image denoising via wavelet thresholding: Orthogonal and biorthogonal, scalar and multiple wavelet transforms," *Imperial College, Statistics Section, Technical Report* TR-98-01, 1998
18. P.J.M. van Laarhoven and E.H.L. Aarts, *Simulated Annealing: Theory and Applications*, Kluwer Academic Publishers, 1987.
19. J.S. Geronimo, D.P. Hardin, and P.R. Massopust, "Fractal functions and wavelet expansions based on several functions," *J. Approx. Theory*, **78(3)**, pp. 373-401, 1994.
20. C.K. Chui and J.A. Lian, "A study of orthonormal multiwavelets," *Appl. Numer. Math.*, vol. 20, pp. 273-298, 1995.
21. L.-X. Shen, H.H. Tan, and J.Y. Tham, "Symmetric-antisymmetric orthonormal multiwavelets and related scalar wavelets," *Applied and Computational Harmonic Analysis (ACHA)*, **8(3)**, pp. 258-279, May 2000.
22. V. Strela, "A note on construction of biorthogonal multi-scaling functions," in *Contemporary Mathematics*, A. Aldroubi and E. B. Lin (eds.), pp.149-157, AMS, 1998.
23. R. Turcajova, "Hermite spline multiwavelets for image modeling," Wavelet applications V, Orlando, FL, *SPIE Proc.*, **3391**, pp. 46-56, April 1998.
24. I.W. Selesnick, "Interpolating multiwavelet bases and the sampling theorem," *IEEE Trans. Signal Processing*, **47(6)**, pp. 1615-1621, June 1999.
25. I.W. Selesnick, "Cardinal multiwavelets and the sampling theorem," In *Proc.* of *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, **3**, pp. 1209 -1212, 1999.
26. J. Lebrun, M. Vetterli, "Balanced multiwavelets Theory and design," *IEEE Trans. Signal Processing*, **46(4)**, pp.1119-1124, April 1998.