CURVELET-BASED TEXTURE CLASSIFICATION IN COMPUTERIZED CRITICAL GLEASON GRADING OF PROSTATE CANCER HISTOLOGICAL IMAGES

by

Wen-Chyi Lin

BS, Chung-Cheng Institute of Technology, 1998

MS, University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

The Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Wen-Chyi Lin

It was defended on

July 14, 2017

and approved by

Ching-Chung Li, PhD, Professor,
Department of Electrical and Computer Engineering

Luis F. Chaparro, PhD, Associate Professor,
Department of Electrical and Computer Engineering

Zhi-Hong Mao, PhD, Associate Professor,
Department of Electrical and Computer Engineering

Shi-Kuo Chang, PhD, Professor, Department of Computer Science

Mingui Sun, PhD, Professor of Neurological Surgery,
Departments of Bioengineering and Electrical Engineering

Dissertation Director: Ching-Chung Li, PhD, Professor,
Department of Electrical and Computer Engineering

CURVELET-BASED TEXTURE CLASSIFICATION IN COMPUTERIZED CRITICAL GLEASON GRADING OF PROSTATE CANCER HISTOLOGICAL IMAGES

Wen-Chyi Lin, PhD

University of Pittsburgh, 2017

Classical multi-resolution image processing using wavelets provides an efficient analysis of image characteristics represented in terms of pixel-based singularities such as connected edge pixels of objects and texture elements given by the pixel intensity statistics. Curvelet transform is a recently developed approach based on curved singularities that provides a more sparse representation for a variety of directional multi-resolution image processing tasks such as denoising and texture analysis. The objective of this research is to develop a multi-class classifier for the automated classification of Gleason patterns of prostate cancer histological images with the utilization of curvelet-based texture analysis. This problem of computer-aided recognition of four pattern classes between Gleason Score 6 (primary Gleason grade 3 plus secondary Gleason grade 3) and Gleason Score 8 (both primary and secondary grades 4) is of critical importance affecting treatment decision and patients' quality of life. Multiple spatial sampling within each histological image is examined through the curvelet transform, the significant curvelet coefficient at each location of an image patch is obtained by maximization with respect to all curvelet orientations at a given location which represents the apparent curved-based singularity such as a short edge segment in the image structure. This sparser representation reduces greatly the redundancy in the original set of curvelet coefficients. The statistical textural features

are extracted from these curvelet coefficients at multiple scales. We have designed a 2-level 4-class classification scheme, attempting to mimic the human expert's decision process. It consists of two Gaussian kernel support vector machines, one support vector machine in each level and each is incorporated with a voting mechanism from classifications of multiple windowed patches in an image to reach the final decision for the image. At level 1, the support vector machine with voting is trained to ascertain the classification of Gleason grade 3 and grade 4, thus Gleason score 6 and score 8, by unanimous votes to one of the two classes, while the mixture voting inside the margin between decision boundaries will be assigned to the third class for consideration at level 2. The support vector machine in level 2 with supplemental features is trained to classify an image patch to Gleason grade 3+4 or 4+3 and the majority decision from multiple patches to consolidate the two-class discrimination of the image within Gleason score 7, or else, assign to an Indecision category. The developed tree classifier with voting from sampled image patches is distinct from the traditional voting by multiple machines. With a database of TMA prostate histological images from Urology/Pathology Laboratory of the Johns Hopkins Medical Center, the classifier using curvelet-based statistical texture features for recognition of 4-class critical Gleason scores was successfully trained and tested achieving a remarkable performance with 97.91% overall 4-class validation accuracy and 95.83% testing accuracy. This lends to an expectation of more testing and further improvement toward a plausible practical implementation.

Index Terms— Curvelets, local maximum curvelet coefficient, tissue texture classification, prostate cancer, Gleason grading, Gleason scores

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation is the outcome of a wonderful four-year working experience under the guidance of my advisor Professor C. C. Li, to whom I wish to express my gratitude and appreciation. His vision, creativeness, and enthusiasm have fueled this research and have a lasting influence in my career as a role model. I would like to thank Professor Luis F. Chaparro, Professor Zhi-Hong Mao, Professor Shi-Kuo Chang, and Professor Mingui Sun, my committee members, for their helpful reviews and comments. I would also like to acknowledge Dr. Robert W. Veltri and Dr. Jonathan I. Epstein of the Departments of Urology and Oncology at the Johns Hopkins University, for their valuable advice, generous help, fruitful collaborations and providing the Tissue Microarray dataset of prostate cancer histological images used in this research. This work was supported by a PhD Fellowship from the Taiwan R.O.C. government, which I gratefully acknowledge.

1.0   INTRODUCTION

Gleason grading system has been established as the standard for interpreting prostate carcinoma by expert pathologists based on microscopic tissue images (TMA) from needle biopsies [49-50]. Gleason grade is categorized into 1 to 5, and increases based on the cumulative loss of regular glandular structure which reflects the increasing degree of malignancy aggressive phenotype. The major work of computer-aided classification as reported in the published literatures is mainly focused on the classification between primary grade 3, 4 and 5 by microstructure, morphology or texture methods [56, 58, 109]. Approaches based on architecture require tissue segmentation before performing feature extraction and thus rely heavily on correct identification of prostate nuclei and glandular structures as shown in Figure 1. Texture-based features are used broadly in data driven methods because of the simplicity of feature extraction and effectiveness of representing histopathology characteristics displayed in images.

As compared to the structure-based and segmentation oriented techniques [58, 59, 111], the texture features extracted from wavelet and multiwavelet coefficients have been shown to be able to achieve more preferable results [4, 53]. Moreover, the wavelet transform is capable of capturing texture complexity that provided by spatial and frequency information across several scales with multiresolution analysis. Nevertheless, the wavelet-based methods share a

disadvantage of extracting features from transform coefficients which are limited in directional selectivity and capture only point singularities.
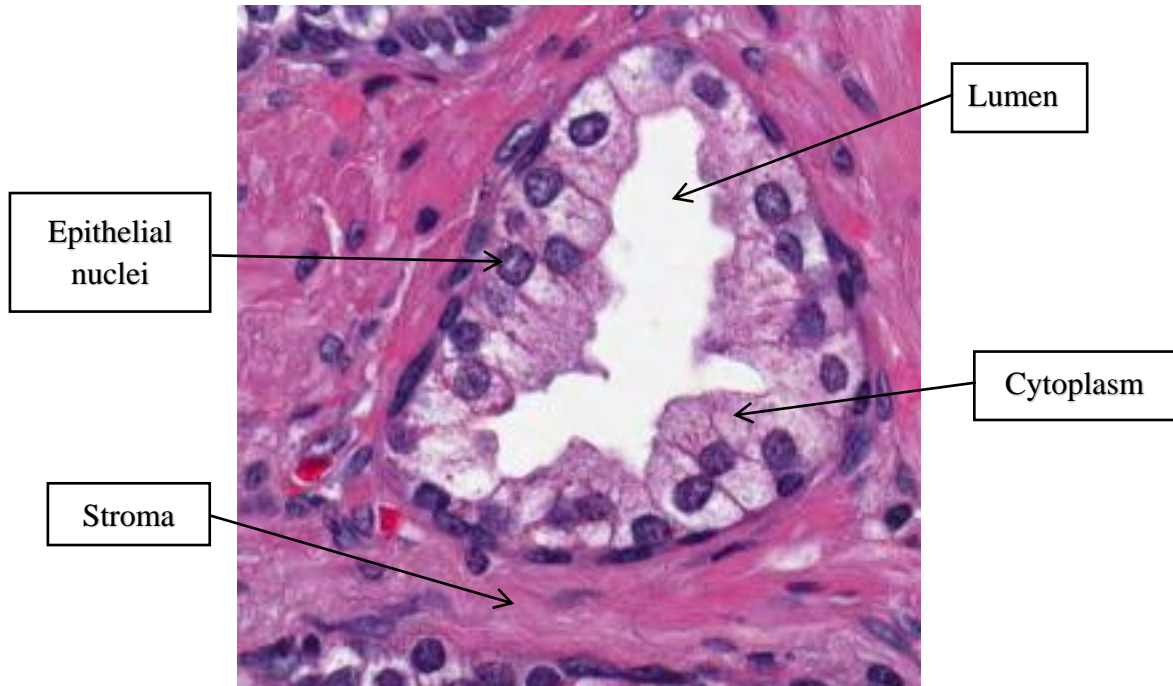
Figure 1. Pattern and structure of a benign gland

Gleason score (GS) is the summation of the primary grade and the secondary grade, ranging from 5 to 10 where the reliable recognition between GS 3+3, 3+4, 4+3 and 4+4 is of crucial importance. However, most published studies dealt with primary Gleason grade 3, 4 and 5 while the classification for critical Gleason patterns, the GS 3+3, 3+4, 4+3 and 4+4, is less reported. The intermediate-grade carcinoma is the mid-point between the low-grade and high-grade and it receives the most lack of consensus in second-opinion assessments which is considered as the critical Gleason patterns where the inaccurate determination of prognosis in terms of over- or under-assessment will greatly affect appropriate treatment and inevitably result in undesirable quality of life for patients [38-39, 70].

With regard to the advances of image processing and classification methods, the second generation curvelet transform developed by Candes and Donoho [11-13] is a generalization of wavelet transform for optimal sparse representation of a class of continuous functions with curvilinear singularity, i.e., discontinuity along a curve with bounded curvature. Although the fundamental of curvelet transform has been introduced to application fields [5, 18, 21], the utilization of all available curvelet coefficients showed a high redundancy ratio that needs to be deliberated in extraction of discriminant information. The machine for classification of patterns of intermediate class, such as Gleason patterns 3+4 and 4+3 in prostate cancer images, which are correlated to other two classes such as class of Gleason patters 3+3 and of Gleason patters 4+4, needs to be meticulously designed to achieve good accuracy.

The objective of this research is, first, to explore the utilization of the most significant curvelet coefficients for concise and realistic representation of image characteristics, based on which to extract discriminating texture features of prostate cancer tissue images, and second, to design a classifier for computer-aided classification of 4 classes of critical Gleason patterns where two classes are correlated with the other two classes. The selected curvelet-based texture features will be used; along with the design strategy, to achieve a performance of high and reliable classification accuracy.

The thesis is organized as follows. First, a brief review of curvelet transform in the context of image processing is given in Chapter 2. Chapter 3 presents the utilization of maximum curvelet coefficients to extract statistical texture features from prostate histological images of critical Gleason patterns. Our study on the classifier design for computer-aided classification of four classes of critical Gleason patterns of prostate cancer tissue images and our experimental

results are presented in Chapter 4. Chapter 5 summarizes the major contributions of the thesis and

gives suggestions for future research.

2.0   THE METHOD OF CURVELET TRANSFORM

Wavelet transform and its generalizations have become the fundamental approach for multi-resolution analysis. Multiresolution image analysis [23] projects an image into spaces in multiple scales (resolutions) which allows us to focus on characteristics pertaining to specific scales, the composition of which makes an efficient representation for analysis, measurement and storage. A wavelet is a basis function as a building block for functional representation in multiple scales. Wavelet transform provides a sparse representation for functions with point singularities; a point singularity means that there is a discontinuity or significant change in the immediate neighborhood of a point, such as an edge pixel in a 2-dimensional image. Edge pixels need to be appropriately connected to obtain object boundaries in image segmentation, or to form texture elements or textons in micro and macro scales, or in a mixture manner, for texture extraction. Well-known successful results have been obtained for image/video compression, denoising, image segmentation, image texture classification, and image fusion, etc [5, 22, 24, 25, 80].

In general, a 2-dimensional wavelet is expressed as a tensor product of two 1-dimensional wavelets and therefore, it can isolate well a point singularity, i.e., the discontinuity across an edge pixel, effortlessly, but the directional selectivity is limited to only three orientations: horizontal, vertical and diagonal. Ridgelet transform and curvelet transforms are recent extensions of wavelet transform to consider line and curve singularities [26, 27].

2.1   RIDGELET TRANSFORM

An anisotropic geometric wavelet, named ridgelet, was introduced by Candés and Donoho [27] in 1998 to overcome the limitation of the standard 2-D wavelet transform. Ridgelet transform is a type of multiscale direction-selective transform that uses ridgelet as the basis element which has high directional selectivity. The ridgelet transform processes the image data by first performing line integrals over different radial orientations at various locations to acquire the Radon transform [28], and then applying wavelet transform to the Radon coefficient profiles. In this way, the wavelet transform handles the sharp point singularities in the Radon domain which reflect the line singularities in the pertinent image locations. It provides a very effective technique to perform sparse directional analysis [16, 29]. This sparsity property of the ridgelet transform enables better texture discrimination for a class of 2-D images than that of the wavelet transform.

The first generation of curvelet transform is a generalization of the ridgelet transform. As edges in images, in general, appears to be curvilinear rather than straight lines, at sufficiently fine scales, a curved edge, nevertheless, looks almost straight where the ridgelet can be utilized upon the premise. To analyze curve singularities in 2-D images, a natural way is to consider a curve which is piecewise linear, and thus to consider a smooth partition of the image into small blocks band apply the ridgelet transform locally to the partitioned sub-images. This concept of the block ridgelet transform led to the first-generation curvelet transform [29], where curvelet represent curves as a superposition of functions with supports of various lengths and widths obeying the parabolic scaling law width $\approx$ length^2. The processing of the first-generation curvelet transform is given by the following steps:

(1) Subband decomposition (the image is filtered into subbands),

(2) Smooth block partitioning (each subband is smoothly windowed into squares of

appropriate scale,

(3) Renormalization (each resulting square is renormalized to unit scale, and

(4) Ridgelet Analysis (each square is processed with the ridgelet transform).

To obtain high direction sensitivity and anisotropy, the first generation curvelet transform has a redundancy factor of $16J + 1$ whenever J scales are used in subband decomposition [30] but is rather restricted due to the accuracy problem of requiring smoothly partitioned blocks of small size in computing the local ridgelet transform.

2.2   CURVELET TRANSFORM

The second-generation curvelet transform [11] captures the local curved structures with radial 'wedges' constructed by concentric squares in the frequency domain to provide a tight frame for functional representation in the spatial domain with low redundancy. A curvelet is defined inside the wedge and has three parameters: scale, orientation, and spatial location. The curvelet transform is a mapping of a multivariable function into a space spanned by curvelets in multiple scales and multiple orientations. It gives an efficient representation for functions with curved singularities but are smooth away from discontinuities across such curves. In the curvelet domain, the information of prominent edges or curved elements in images is packed into a small number of coefficients, thus the transform yields a very sparse image representation.

A curvelet in the 2-dimensional space is an elementary function $\varphi(x_1, x_2)$ of two spatial variables x_1 and x_2 that is defined primarily over a narrow rectangular support region of short width (e.g., along x_1 axis) and longer length (e.g., along x_2 axis) following the parabolic scaling rule, i.e., the width scaling is equal to the square of the length scaling as illustrated in the right of Figure 2(a)

support for a curvelet in the spatial domain is shown in the right together with its rotated and shifted in the left of the figure. It changes rapidly along x_1 and is smooth along x_2, so its Fourier transform $\hat{\varphi}_j(\omega_1, \omega_2)$ is in a broad high frequency band along ω_1 and is limited to a narrow low frequency band along ω_2, that means, $\hat{\varphi}_j(\omega_1, \omega_2)$ is supported over a narrow wedge sector in the 2-d frequency domain (ω_1, ω_2) where ^ denotes the Fourier transform. When φ is expressed by $\varphi(\rho, \theta)$ in the spatial polar coordinates (ρ, θ), its Fourier transform is given by $\hat{\varphi}_j(r, \theta_t)$ in the frequency polar coordinates (r, θ). If $\varphi(x_1, x_2)$ is rotated by an angle θ_l but without translation, and expressed in the spatial polar coordinates ($\rho, \theta - \theta_l$), its Fourier transform is of the similar pattern appearing along the radial frequency axis r and across a narrow sector of angular frequency θ_t in the polar frequency coordinates ($r, \theta_t - \theta_l$). In Figure 2(b), a polar frequency plane is shown with radial windows in a circular coronae for supporting curvelets of a scale in different orientations. The shaded sector illustrates a narrow radial wedge centered at a long radial frequency line but with a short width in angular frequency compliant with the parabolic scaling rule. It provides the frequency support of the curvelet centered in a specific spatial orientation θ_l. That is how the second generation curvelet is generated.

With both shift (k_1, k_2) in (x_1, x_2) and rotation θ_l, a curvelet at scale j (for $j \geq 0$) is given by

$$\varphi_{j, \theta_l, k}(x_1, x_2) = 2^{3j/4} \varphi(R_{\theta_l}[2^j(x_1 - k_1), \quad 2^{j/2}(x_2 - k_2)]^T)$$

and its Fourier transform is

$$\hat{\varphi}_{j, \theta_l, k}(\omega_1, \omega_2) = 2^{-3j/4} \hat{\varphi}\left(R_{\theta_l}\left[2^{-j}\omega_1, 2^{-j/2}\omega_2\right]^T\right) e^{-i(k_1\omega_1 + k_2\omega_2)} \tag{2.1}$$
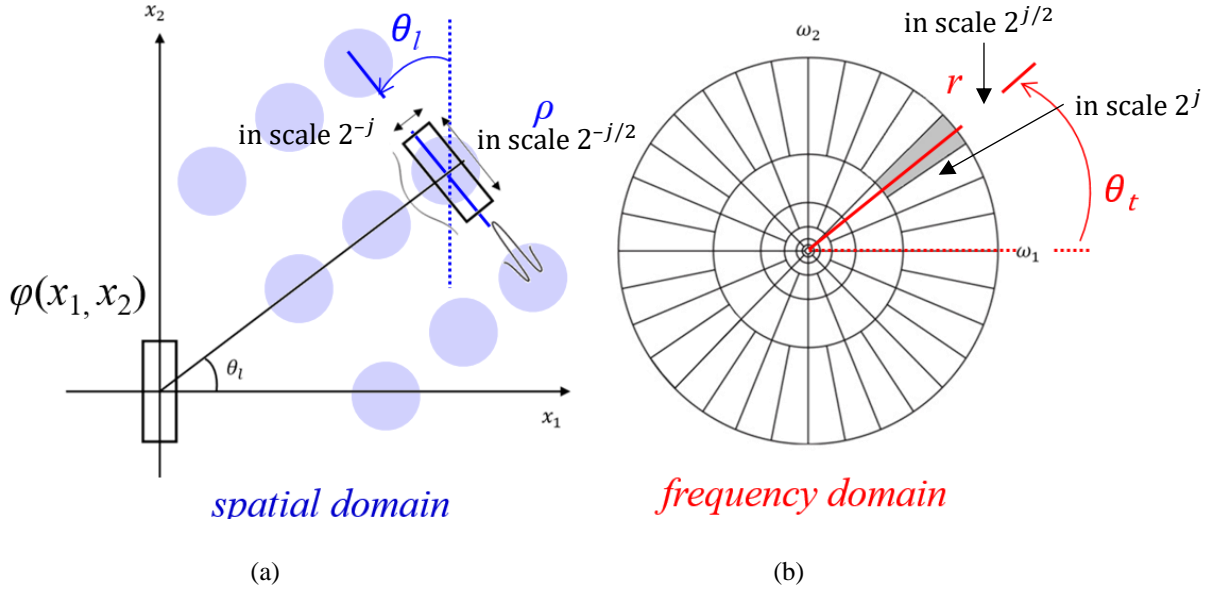
Figure 2. Illustrations of curvelets with parabolic scaling (a) in the spatial domain and (b) in the 2-D frequency plane. Similar curvelet responses can be found in same radial wedges. The shaded sector denotes the frequency support of a shifted and rotated curvelet in (a).

where subscript k denotes a column vector $[k_1, k_2]^T$ and R_{θ_l} is a rotation matrix with respect to the horizontal reference,

$$R_{\theta_l} = \begin{pmatrix} \cos\theta_l & \sin\theta_l \\ -\sin\theta_l & \cos\theta_l \end{pmatrix},$$

The set of $\{\varphi_{j,\theta_l,k}(x_1,x_2)\}$ is a tight frame that can be used to represent a function $f(x_1, x_2)$ by the linear combination of the curvelets $\{\varphi_{j,\theta_l,k}(x_1,x_2)\}$

$$f = \sum_{j,\theta_l,k} c_{j,\theta_l,k}\varphi_{j,\theta_l,k}(x_1,x_2)$$

where the curvelet coefficients $\{c_{j,\theta_l,k}\}$ are given by inner products

$$c_{j,\theta_l,k} = \left\langle f, \varphi_{j,\theta_l,k} \right\rangle \tag{2.2}$$

This gives the curvelet transform.
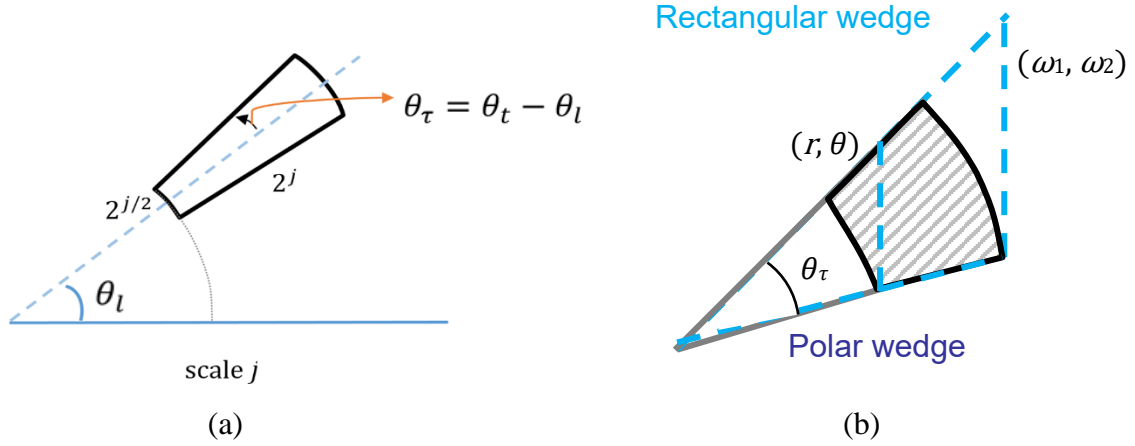
Figure 3. A polar wedge constructed for scale j and orientation θ_l in (a) the frequency domain and in (b) the polar wedge is replaced by the pseudopolar wedge to be adapted to the Cartesian coordinates.

For the discrete curvelet transform, let a frequency domain polar wedge be defined by U_j

$$U_j(r,\theta_\tau) = 2^{\frac{-3j}{4}} W(2^{-j}r)V(2^{\lfloor j/2 \rfloor}\theta_\tau),\tag{2.3}$$

where $W(r)$ and $V(\theta_\tau)$ are a pair of radial window and angular window in the polar coordinates with $r \in (1/2, 2)$,

$$\theta_\tau = q \cdot \frac{\pi}{2} \cdot 2^{-\lfloor j/2 \rfloor} \quad \in (-\frac{\pi}{2} \cdot 2^{-\lfloor j/2 \rfloor}, \ \frac{\pi}{2} \cdot 2^{-\lfloor j/2 \rfloor}), \quad (q = 0, 1, 2, \ldots)$$

$$\theta_\tau = (\theta_t - \theta_l).$$

The $\lfloor j/2 \rfloor$ denotes the integer part of $j/2$. Here r is the normalized radial frequency variable with the normalization constant π and

$$\theta_l = l \cdot (\frac{\pi}{2}) \cdot 2^{-\lfloor j/2 \rfloor}, \ (l = 0, 1, 2, \ldots, N_j - 1),$$

$$\theta_t = t \cdot \frac{\pi}{2} \cdot 2^{-\lfloor j/2 \rfloor}, \qquad \theta_t \in (-\frac{\pi}{2}, \quad \frac{\pi}{2}) \quad (t = 0, 1, 2, \ldots)$$

where N_j denotes the number of polar wedges, thus, the number of orientations considered which is equal to the number of rotation increment over the 2π range. The angular frequency variable $\theta_\tau = q \cdot (\pi/2) \cdot 2^{-\lfloor j/2 \rfloor}$ varies around the orientation θ_l as shown in Figure 3(a). Radial window $W(r)$ and angular window $V(\theta_\tau)$ are smooth non-negative real-valued functions and are subject to the admissibility conditions

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \qquad r \in (\frac{3}{4}, \quad \frac{3}{2}), \tag{2.4}$$

$$\sum_{\ell=0}^{N_j-1} V^2(\theta_t - \theta_l) = 1, \qquad \theta_t \in [0, \quad 2\pi). \tag{2.5}$$

The window V has its width inversely related to the length of the window W, thus, at scale j, $V_j(\theta_\tau)$ is given by the expression V in (2.3) which refers to the shaded region in Fig. 2(b). With the symmetry property of the Fourier transform, the range of θ_t is $(-\pi/2, \pi/2)$. Let the curvelet at scale j without shift be defined in the polar wedge U_j

$$\hat{\varphi}_{j,l}(\omega_1, \omega_2) = U_j(R_{\theta_l}\omega) = U_j(r, \theta_t - \theta_l). \tag{2.6}$$

With a spatial translation (k_1, k_2), it will then be

$$\hat{\varphi}_{j,l,k}(\omega_1, \omega_2) = U_j(R_{\theta_l}\omega)e^{-i(k_1\omega_1 + k_2\omega_2)}, \tag{2.7}$$

its inverse Fourier transform gives

$$\varphi_{j,\theta_l,k}(x_1, x_2).$$

11

Through the Plancherel's theorem, the curvelet coefficients are also given by the inner products in the frequency domain

$$c(j,l,k) = \left\langle f, \varphi_{j,\theta_l,k} \right\rangle := \frac{1}{(2\pi)^2} \left\langle \hat{f}, \hat{\varphi}_{j,\theta_l,k} \right\rangle = \frac{1}{(2\pi)^2} \int \hat{f}(\omega)\overline{\hat{\varphi}_{j,\theta_l,k}(\omega)}d\omega$$

$$= \frac{1}{(2\pi)^2} \int \hat{f}(\omega_1,\omega_2) U_j(R_{\theta_l}\omega)e^{i(k_1\omega_1+k_2\omega_2)}d\omega_1 d\omega_2 \qquad (2.8)$$

The magnitude of curvelet coefficient $c(j,l,k)$ denotes the strength of the curved singularity, a short edge segment at scale j centered at the location $k = (k_1, k_2)$ and in orientation θ_l. In digital implementation, the polar wedge is replaced by the pseudopolar wedge in the Cartesian coordinates as shown in Figure 3(b).

2.2.1 Multiresolution Decomposition and Subband Index

In the above discussion, the scale index $j \geq 0$ is considered, increasing refers to finer resolution. In image processing applications, a given image of the size $2^M \times 2^M$ is in the finest resolution which is designated as the reference scale $j = 0$. Dyadic decomposition into coarser scales $j = -1, -2, \ldots,$ leads to subband images of $2^{M-1} \times 2^{M-1}$, etc. in successively lower frequency band. Consider a scan line of 2^M pixels, its Fourier transform FFT is given by 2^M frequency increments which is dyadically decomposed into M subband with subband index i corresponding to the decomposition scale level j, $(j < 0)$, $i = M + 1 + j$, $(i = 1, 2, \ldots, M)$ as shown in Figure 4. For a radical scan line, this also holds true in terms of radical frequency subband. Referring to Figure 2(a), the vertically orientated narrow rectangle for $j > 0$ will become horizontally orientated for $j < 0$ as illustrated in Figure 5. As j goes to be very negative for coarse scales, the rectangle orientation becomes not

meaningful any more for an oriented curvelet to be defined there. So, the three subbands covering

the lowest frequencies are left out to be contained in a scaling component $\hat{\phi}_o(\omega)$ with

multiresolution analysis, and we relabel the cuvelet subband index by J.

$$J = i - 3 = M + 1 + j - 3 = M - 2 + j. \quad (J = 1, 2, \ldots, M - 3)$$

where J subbands are used in our curvelet transform analysis.
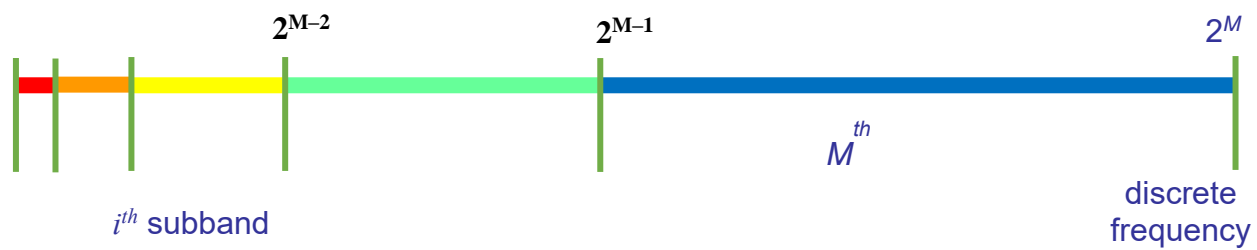
Figure 4. A scan line of 2^M points FFT is divided into M subband
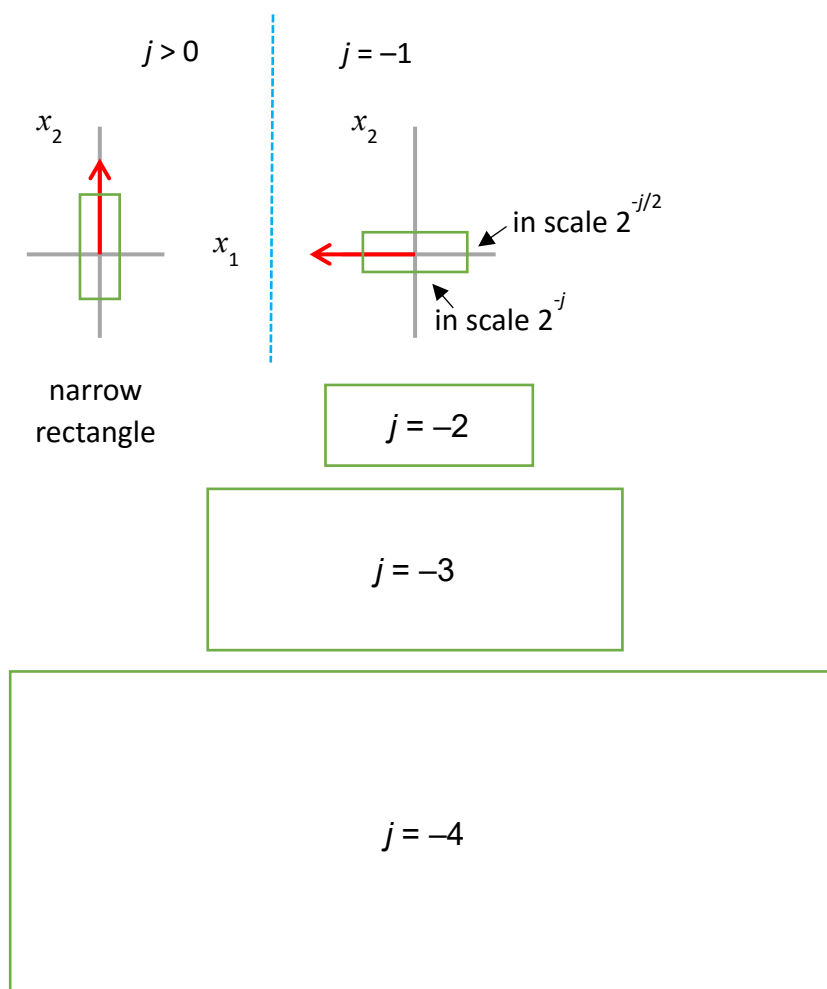
Figure 5. The orientation of the rectangular region changes from vertical to horizontal as the image is

decomposed into coarse scales where $j<0$.

The curvelet transform takes the advantage of the inter-scale orthogonality, the multiscale separation of subband frequency components can be achieved through the 2-D Meyer wavelet analysis [5, 81]. The Meyer wavelet in the frequency domain forms an orthogonal basis with symmetric band-limited functions. The wavelet and scaling functions in the frequency domain are smooth and have compact supports, in the spatial domain they are infinitely differentiable with infinite support. The Meyer wavelet in the Fourier domain is given by

$$
\hat{\psi}(\omega) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} e^{i\omega/2} \sin(\dfrac{\pi}{2} v(\dfrac{3|\omega|}{2\pi} - 1)), & \text{if} & \dfrac{2\pi}{3} \le |\omega| \le \dfrac{4\pi}{3} \\[2mm] \dfrac{1}{\sqrt{2\pi}} e^{i\omega/2} \cos(\dfrac{\pi}{2} v(\dfrac{3|\omega|}{4\pi} - 1)), & \text{if} & \dfrac{4\pi}{3} \le |\omega| \le \dfrac{8\pi}{3} \\[2mm] 0 & & elsewhere \end{cases} \tag{2.9}
$$

Where the auxiliary function $v(\Delta)$ is a smooth non-negative function goes from 0 to 1 on the interval [0, 1] and satisfying

$$
v(\Delta) = \begin{cases} 0 & \text{if} & \Delta \le 0 \\ 1 & \text{if} & \Delta \ge 1 \end{cases} \tag{2.10}
$$

and

$$
v(\Delta) + v(1 - \Delta) = 1. \tag{2.11}
$$

a classical auxiliary function is:

$$
v(x) = x^4 (35 - 84x + 70x^2 - 20x^3). \tag{2.12}
$$

which given a smooth transition v as Δ goes from 0 to 1 as shown in Figure 6.
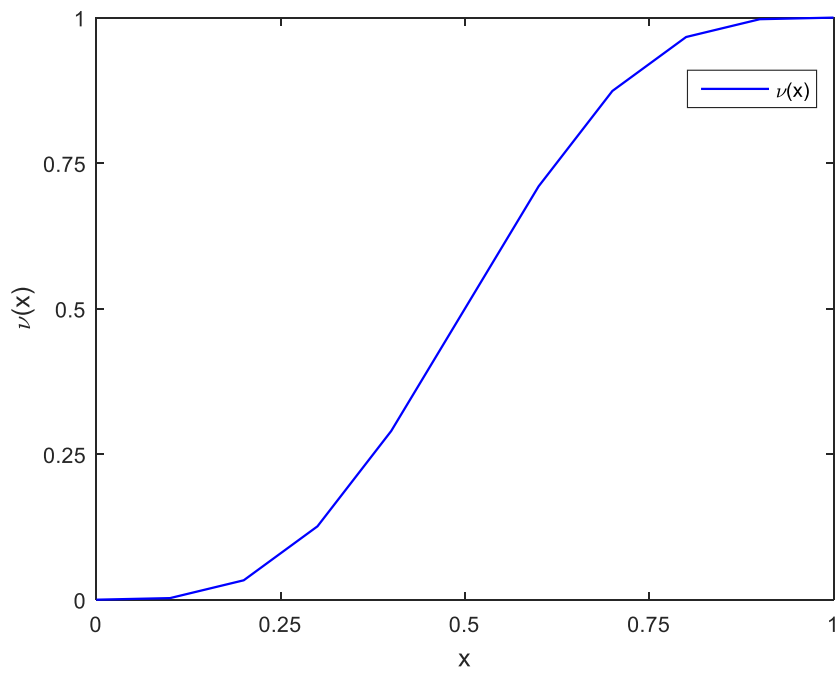
(a)
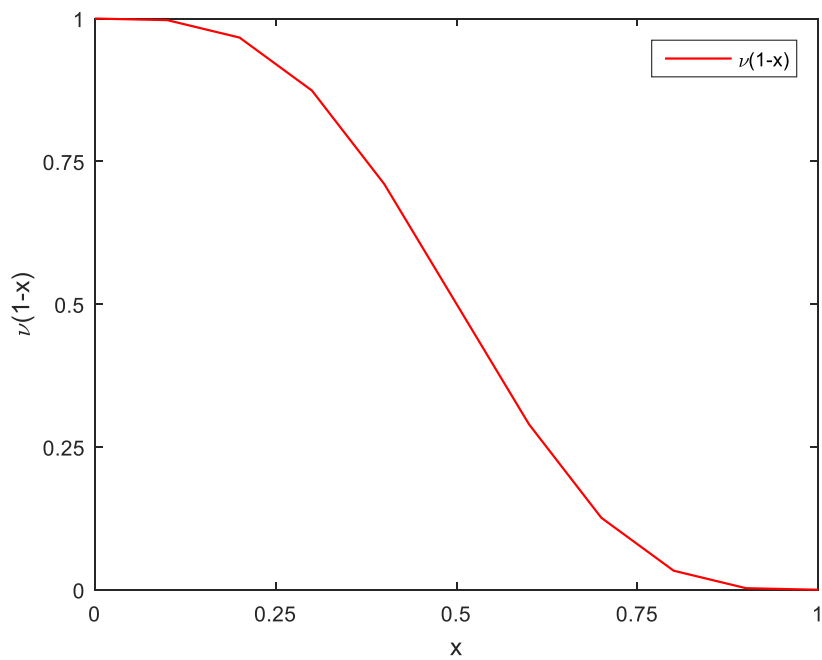
(b)

Figure 6. The auxiliary function (a) $v(x)$ and (b) $v(1–x)$.

The Meyer scaling function in the frequency domain is

$$\hat{\phi}(\omega) = \begin{cases} \dfrac{1}{\sqrt{2\pi}}, & \text{if} & |\omega| \le \dfrac{2\pi}{3} \\[3mm] \dfrac{1}{\sqrt{2\pi}} \cos(\dfrac{\pi}{2} v(\dfrac{3|\omega|}{2\pi} - 1)), & \text{if} & \dfrac{2\pi}{3} \le |\omega| \le \dfrac{4\pi}{3} \\[3mm] 0 & \text{otherwise} \end{cases} \qquad (2.13)$$

Figure 7 displays the graphs of the Fourier transform of the Meyer wavelet and scaling functions as well as the functions in the time domain.
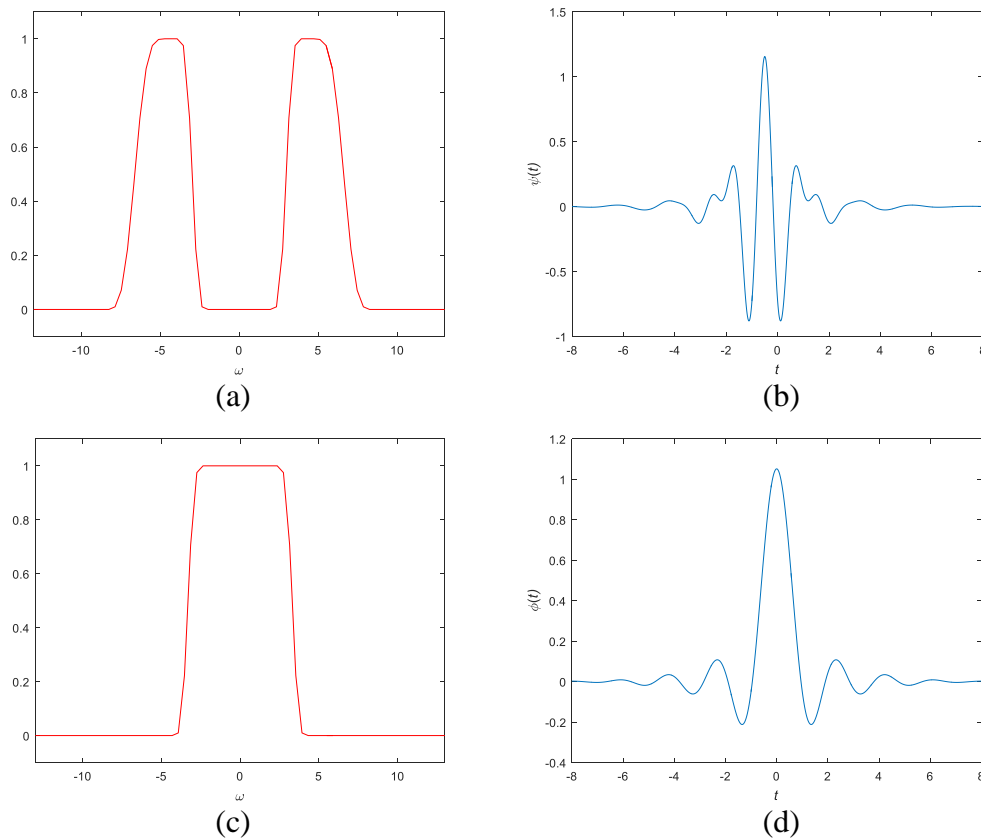
Figure 7. Meyer wavelet and scaling functions (a)(c) in the frequency domain and (b)(d) in the spatial domain.

17

The Meyer wavelet function $\hat{\psi}(\omega)$ can be obtained by taking the product of the dilated Meyer scaling function $\hat{\phi}(\omega/2)$ and a 2π periodic function m_0 in $L^2([0, 2\pi])$ constructed from $\hat{\phi}$ [82], as shown in Figure 8.
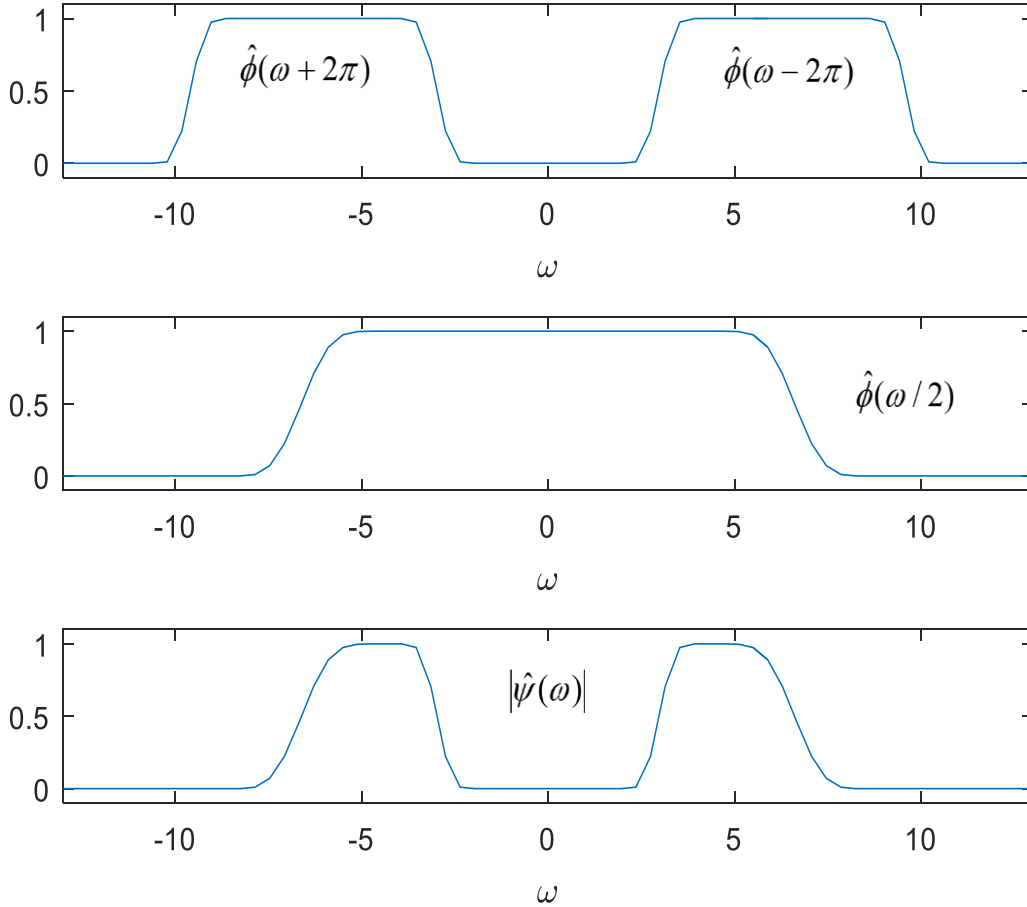
Figure 8. Construction of Meyer wavelet function in the frequency domain

In the similar manner, the radial frequency concentric window $W_j(\omega)$ used in Eq. (2.3) and Eq. (2.6) in constructing curvelet $\hat{\phi}_{j,l}(\omega)$ may be obtained by multiplying the dilated Meyer scaling function along the radial frequency with a 2π-periodic shifted function at scale j and subtracting its next coarse version. Finally leading to

$$\Phi_0(\omega)^2 + \sum_{j \geq 0} \tilde{W}_j^2(\omega) = 1. \tag{2.14}$$

18

where ω is the radial frequency *r* used earlier. The angular window $V_j(\omega_\theta) = V(2^{\lfloor j/2 \rfloor}\omega_\theta)$ can be built with the Meyer scaling window [20] to achieve a smooth separation of angular frequency components $\omega_\theta = \theta_\tau$ in the angular direction. The angular window size decreases as the scales become finer (*j* increases). With the radial frequency *r* and the angular frequency ω_θ, the wedge window is then defined as

$$\tilde{U}_j(\omega) := \tilde{W}_j(r)V_j(\omega_\theta).$$ (2.15)

Introduce a set of slopes indicating the orientation increment $\tan\theta_\tau := q \cdot 2^{-\lfloor j/2 \rfloor}$, $q = -2^{\lfloor j/2 \rfloor}, \cdots, 2^{\lfloor j/2 \rfloor} - 1$, and a shear matrix in the frequency domain

$$S_\theta := \begin{pmatrix} 1 & 0 \\ -\tan\theta & 1 \end{pmatrix}$$

the angular wedge sector and, thus, the curvelet $\hat{\phi}_{j,l}(\omega)$ is given by

$$\tilde{U}_{j,l}(\omega) := W_j(\omega)V_j(S_{\theta_\tau}\omega).$$ (2.16)

The scheme of the curvelet transform in the frequency domain is shown in Figure 9 which displays a 5-scale curvelet decomposition of a $2^8 \times 2^8$ image where the number of angular of partitions at the given scale (subband 3-4) is set to 16. The top portion of Figure 10 demonstrates curvelets in the frequency domain at five subbands in four different orientations, the bottom portion shows their corresponding 1-D Meyer wavelet and scaling functions along the frequency axis. The support of the highest scale Meyer wavelet function (subband 5) exceeds the half sampling frequency 1 (normalized by π) and thus, is periodically extended to 4/3 as shown in the red dotted line.
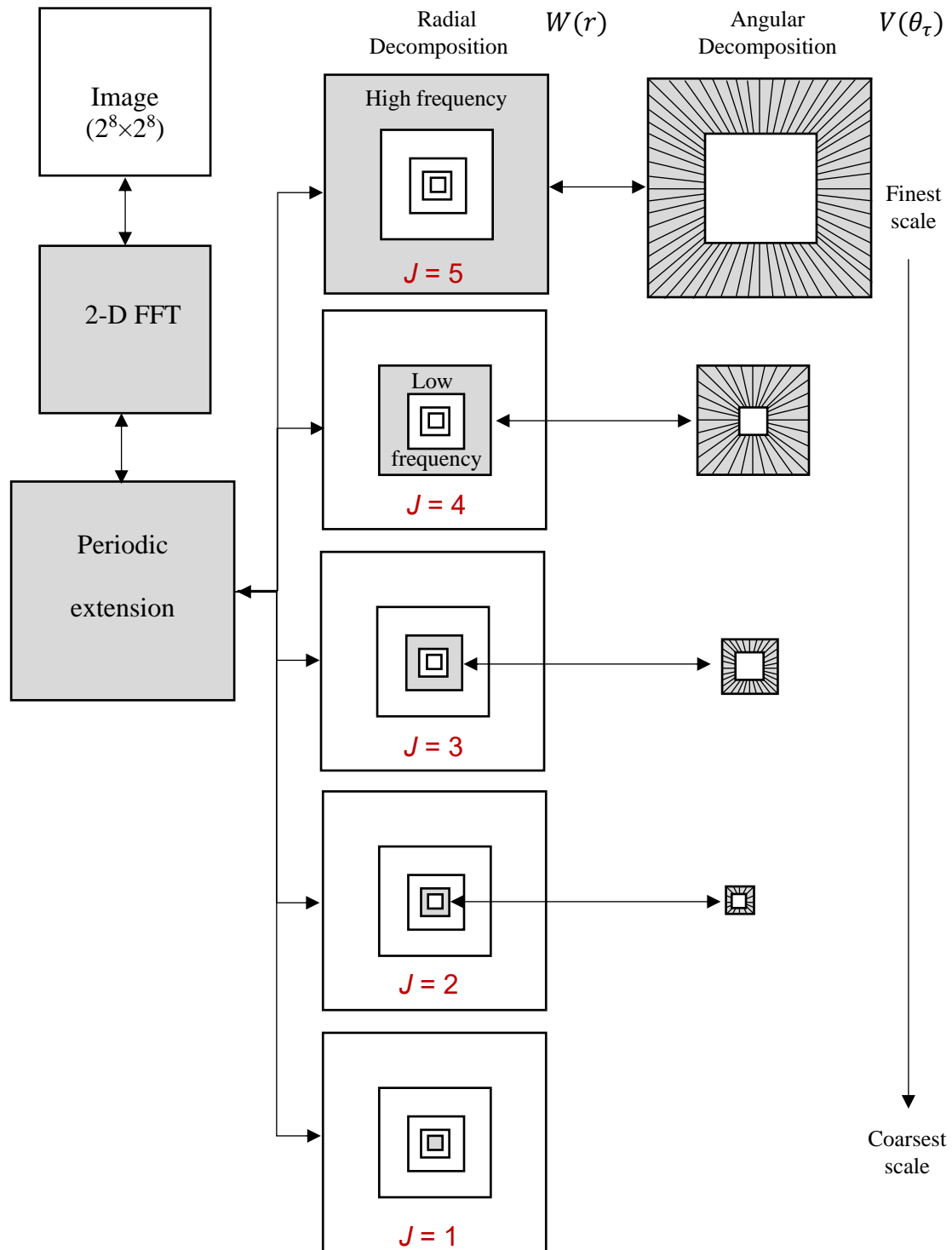
Figure 9. Schematic structure of the fast discrete curvelet transform. Here shows a 5-scale decomposition and the number of angles of next scale is set to 16. The double arrow signs denote that the forward transform is invertible.
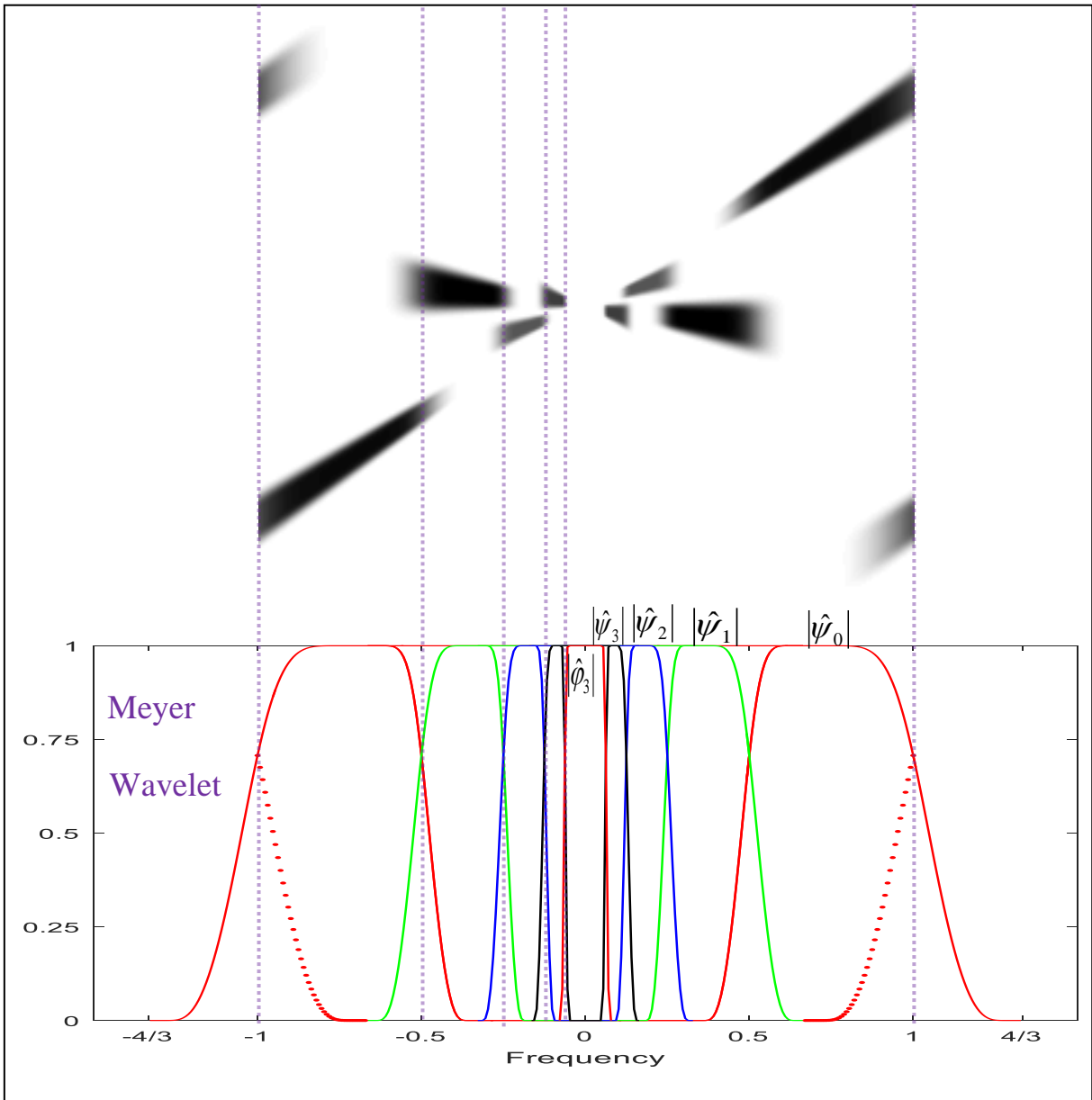
Figure 10. Curvelets in the radial frequency domain at five subbands in different orientations (top). Their corresponding 1-D Meyer wavelet functions and scaling function.

Since curvelet coefficients and curvelets are computed by the inverse Fourier transform via a wedge sector $\tilde{U}_{j,l}(\omega)$ given in Eq. (2.16), the smoothness near the boundaries of the compact W and V windows nevertheless introduces unwanted fluctuations in the spatial domain. This is shown in Figure 11 (a) and (b).

(a)

(b)

(c)

(d)
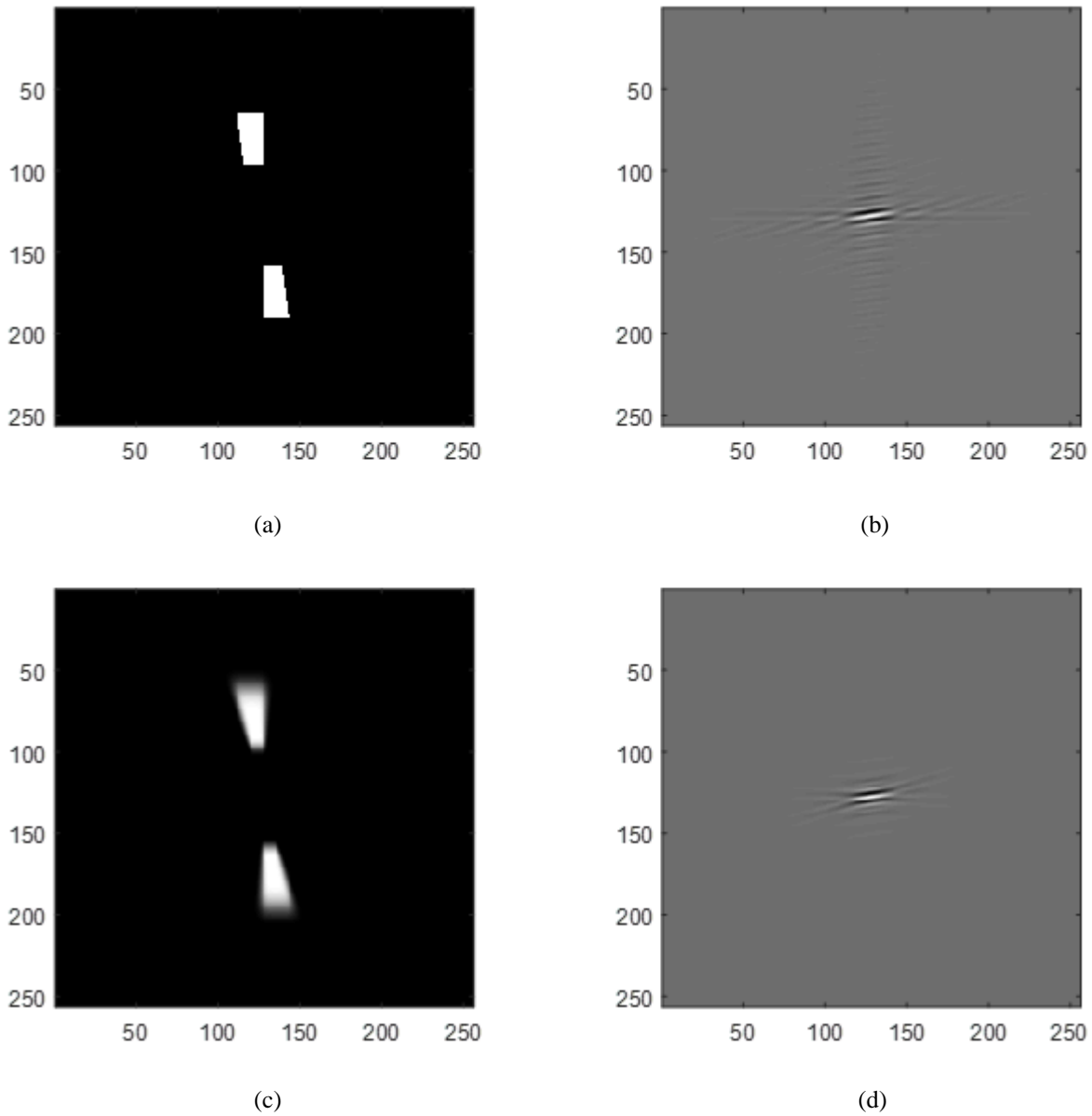
Figure 11. Comparison of the curvelets obtained from the rectangle-edge wedge (a)(b) and the smooth wedge (c)(d). The curvelet has a fast decay with a smooth frequency support near the edges of the wedge.

2.2.2 Computation of Discrete Curvelet Transform Via Frequency Wrapping
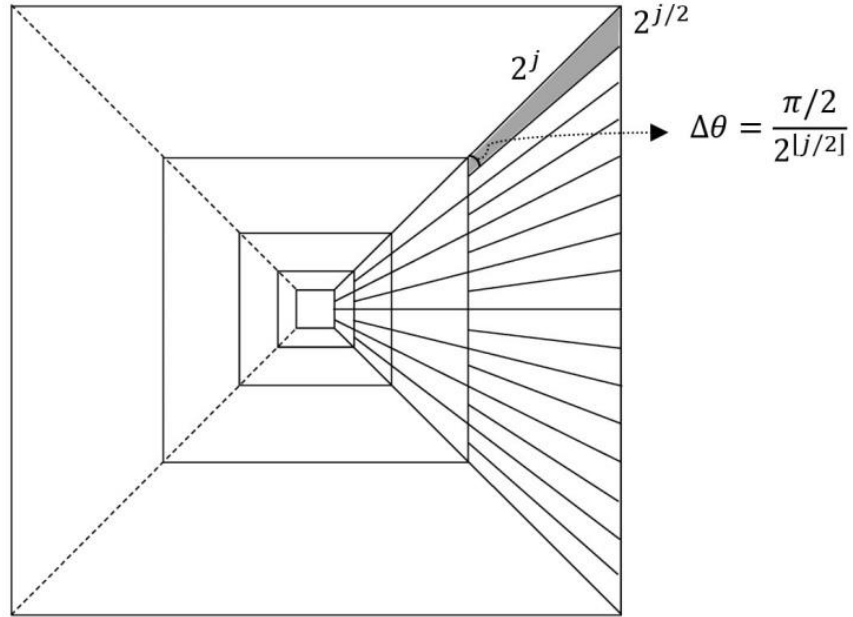
Figure 12. The digital coronae in the frequency plane with pseudo-polar coordinates, trapezoidal wedges are shown also satisfying parabolic scaling rule.

The discrete curvelet transform can be obtained through the inner product in the frequency domain as indicated in Eq. (2.8) and in Figures 2(b) and 3(a) where, for one scale j, the curvelet functions with different orientations are well tiled in a circular shell or coronae. Since the FFT of an image is in the rectangular coordinates, the wedges need to be extend to trapezoidal wedge tiled as concentric rectangular shells as illustrated in Figure 12 to accommodate for computation in rectangular frequency coordinates. The wedges are then in different trapezoidal shapes and incremental orientations of successive wedges are not uniform. Special care must be taken to facilitate the computation. There are two different algorithms for computing the fast digital curvelet transform developed by Candés and colleagues [12, 15]: one is called the unequispaced

FFT approach, and the other is called frequency wrapping approach which is the one we use. The frequency wrapping approach may be briefly explained by the sketch in Figure 13 where a shaded trapezoidal wedge at scale j is under a shearing process, is mapped to a parallel-pipe shaped support as shown in (a). The tiling of the parallel pipes, shown in Figure 13(b), which is geometrically periodic in either vertical or horizontal direction and each contains the identical trapezoidal information will enable a wrapping process by mapping into a rectangular region centered at the origin of the frequency plane where the data samples in trapezoidal wedges in two neighboring parallel pipes are mapped into the rectangular wedge as illustrated by the shaded parts enclosed in the rectangle. The frequency information contained in the broken pieces in the rectangular wedge is the same as in the parallel pipe and, thus, in the original trapezoidal wedge, except they are re-indexed components of the original data. Use this to compute the inner product with the given image to obtain the same inner product. In this way, the inner product can be computed for each wedge and immediately followed by the inverse FFT to obtain the contribution to the curvelet coefficients from that original wedge with the trapezoidal support. Software for both algorithms are freely available in Candés' laboratory [15], we have used the second algorithm in our study of the curvelet-based texture feature analysis.

(a)   (b)
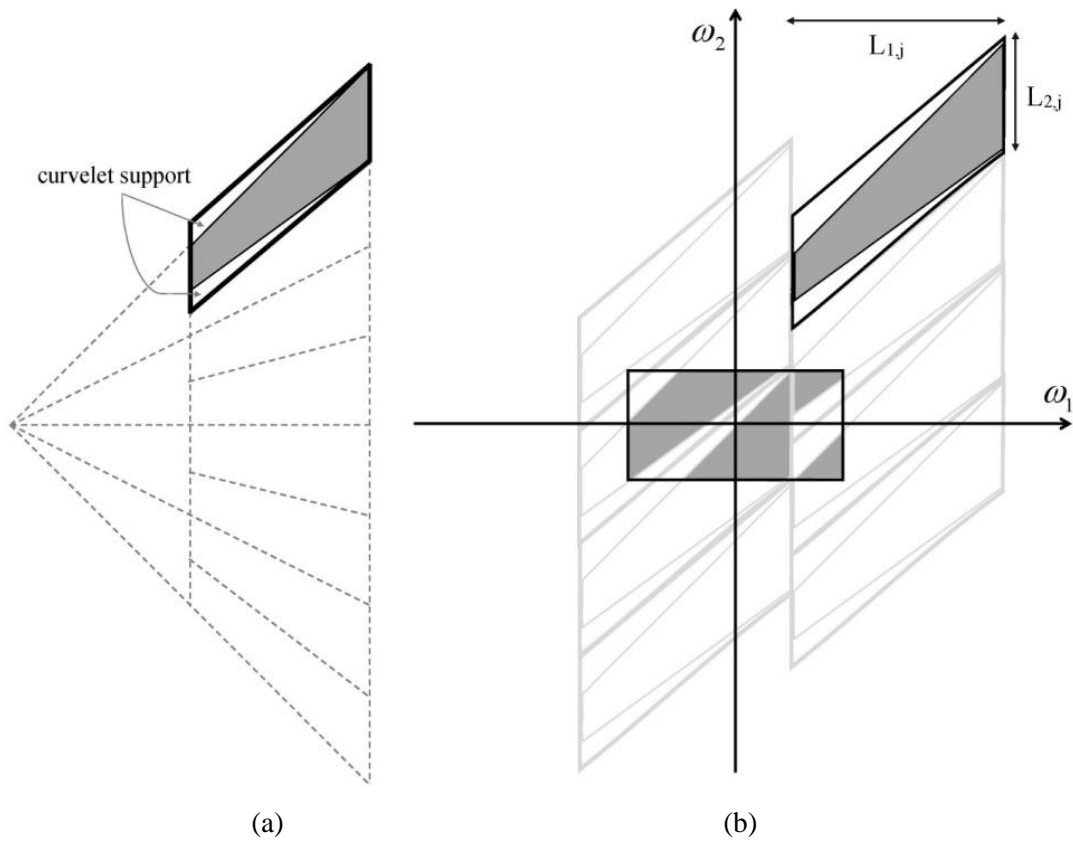
Figure 13. The schematic diagram to illuminate the concept of the wrapping algorithm for computing

digital curvelet coefficients at a given scale.

3.0   CURVELET-BASED TEXTURE FEATURES OF CRITICAL PROSTATE CANCER HISTOLOGICAL IMAGES

In this chapter, we investigate the property of discrete curvelet transform coefficients of prostate cancer tissue images that are used to extract texture features. We first examine the property of curvelet coefficients denoting edge curve segments and their distributions. From the data structure of the coefficients, we consider an efficient and effective way using maximum curvelet coefficients to extract the curvelet-based texture information in the sense of sparsity. Such multiscale texture decomposition using the highly anisotropic characteristics of the curvelets are then studied for discrimination of four critical Gleason patterns in the prostate cancer tissue images.

3.1   CURVELET COEFFICIENT DISTRIBUTIONS

Texture evaluation provides quantitative measures for describing image texture content within a region of interest that are used in scene recognition and classification. In the previous chapter, we have reviewed the relationship between the directional wavelet transform and curvelet transform. Although various wavelet-based texture classification methods have shown many successful applications, the assessment on texture descriptions using the curvelet transform, especially on the efficient utilization of curvelet coefficients, remains to be fully developed.

3.1.1 Curvelet Coefficients and Image Edge Segments

The curvelet coefficients $c(j,\theta_l,k)$ in scale j denotes the strength of the discontinuity, a directional feature or an edge segment in orientation θ_l in an image centered at the spatial location $\boldsymbol{k}$. The waveform of a curvelet shown in Figure 14 is a well-defined function that has very large gradient amplitude across its width and is smooth along its length. When it is used as a basic unit in representation of a sharp change in image brightness, its coefficient magnitude is associated with the degree of change in the direction orthogonal to the longitudinal direction of the curvelet which reflects the boundary of a constituent object such as nucleus or glandular cell in the prostate tissue image. The curvelet coefficients in an image capture curve segments of varying strength over all discrete orientations, but only a small percentage of the coefficients are strong and representing the meaningful edges while a large number of weak coefficients simply describe the very blurred boundaries.
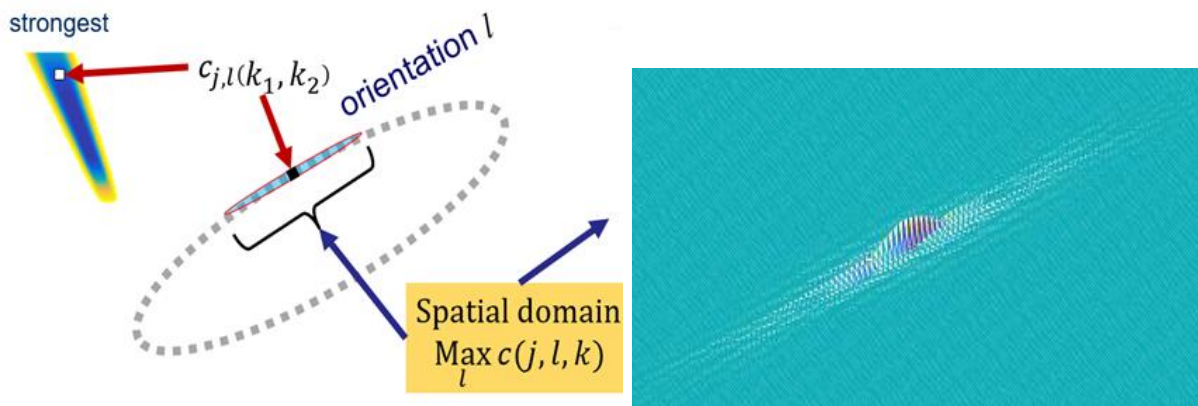
Figure 14. The strongest curvelet coefficient among all orientations indicates the main trace of the line segment in the spatial domain

3.1.2 Curvelet Coefficient Distributions

In image classification problems, it is essential to ascertain that the strong curvelet coefficients would carry useful information of describing the texture content of an image and have enough discriminative power to classify the texture patterns. However, in real situations, most of the curvelet coefficients may have very small magnitudes except those indicating object edges which are much fewer in percentage. As shown in Figure 15, the distributions of curvelet coefficients in two scales of two prostate tissue images of two primary Gleason grades are characterized by high-pitched peaks at zero coefficient magnitude and extended tails at large coefficient magnitudes on both sides (positive and negative). The histograms quantize the curvelet coefficients into limited bins and the small portions of the important edge information are distributed at both ends of the magnitude interval which could make the high order statistical moments play a key role in classification. On the other hand, the generalized Gaussian-like distribution (GGD) contributed by the overwhelming number of coefficients near zero value would not be able to provide the needed discriminative information in the classification task.
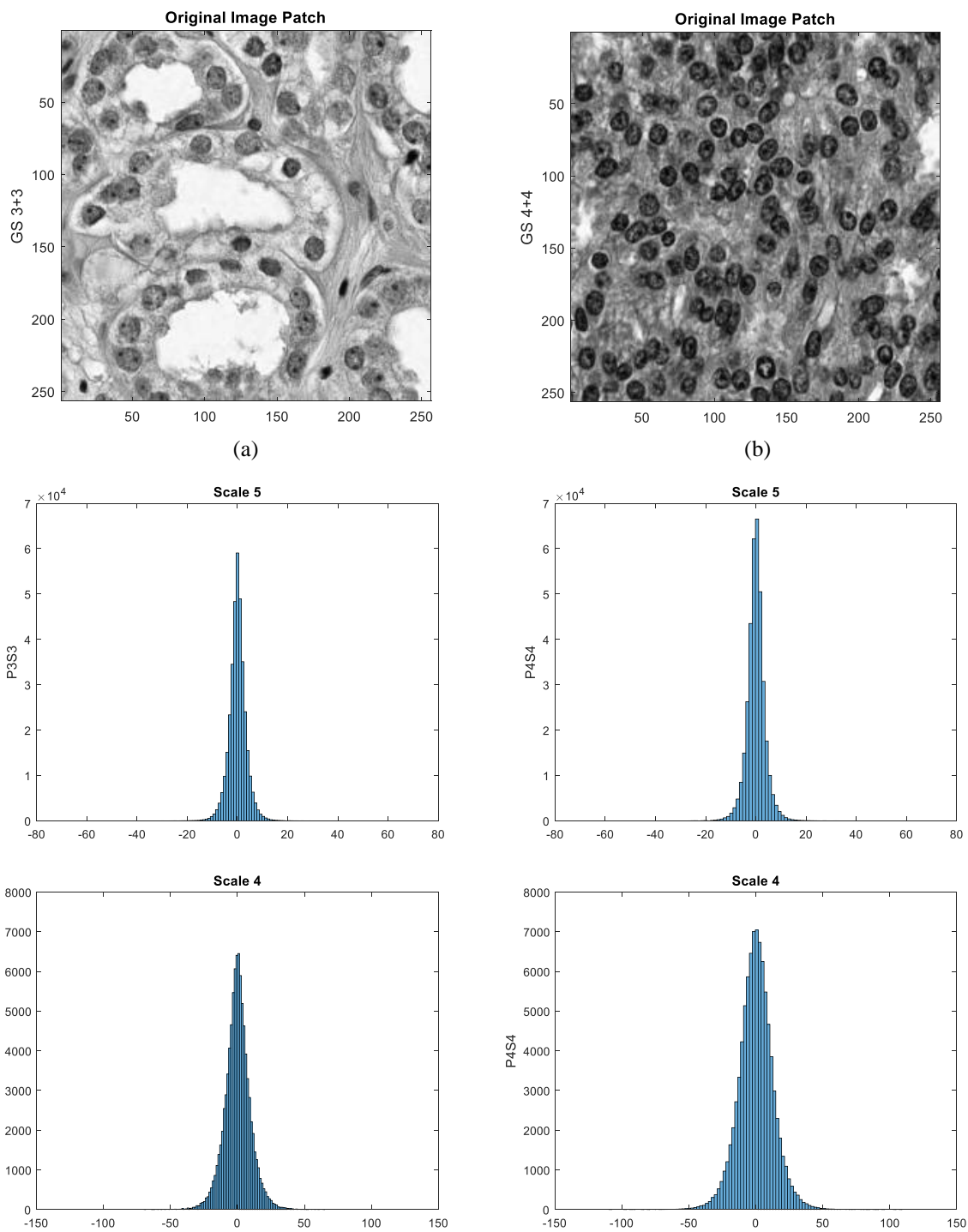
Figure 15. H&E stained prostate histological images (a) of Gleason grade 3 and (b) Gleason grade 4 (top row), histograms of their curvelet coefficients at subband 5 (middle row) and at subband 4 (bottom row).

3.2   MAXIMUM CURVELET COEFFICIENTS

Edge information is very important for image texture analysis because high frequencies in the spectral domain contain more information on texture discrimination properties. The value of curvelet coefficients $c_{j,l,k}$ at position (k_1, k_2) under scale j denotes the strength of the curvelet component oriented at angle θ_l in the representation of an image function $f(x_1, x_2)$. It contains the information on edginess coordinated along a short path of connected edge pixels in that orientation. Naturally, it would be advantageous to extract texture features in the curvelet coefficient space. The tails of the curvelet coefficient distribution are embedded with the edge information but the modeling of coefficient distribution by Generalized Gaussian distribution which led to texture features mainly utilizing the shape and scale parameters is not effective enough for general applications. Further exploration in texture classification is desired.

3.2.1   The Sparsity

Sparsity is a fundamental consideration in modeling physical phenomena leading to sparse segmentation which would permit efficient data processing, e.g., accurate statistical estimation, discriminant feature extraction and classification.

In the discrete implementation of curvelet transform, the Fourier transform of the smooth curvelet widow is multiplied with the Fourier transformed of the image data that is carried out through the wrapping process as we reviewed in 2.2.2. After applying the Inverse Fourier transform to the wrapped FFTs to obtain the curvelet coefficients, the resulting coefficients are saved in the arrays corresponding to the orientations l and according to their relative quadrants.

30

The coefficient arrays are allocated from the $\frac{3}{4}\pi$ in the North quadrant to the East quadrant, South quadrant and the West quadrant. As the consequence of the antisymmetric property of the Fast Fourier transform, the resulting coefficient arrays in the North and the East quadrants will be identical to their South and West counterparts. Therefore, in order to achieve real-valued curvelet coefficients, the real part and the imaginary part of the computed coefficients can be regarded as the cosine curvelet and sine curvelet [15] which are stockpiled in a pair of symmetry arrays as shown in Figure 16 where the cosine curvelet and the sine curvelet of the initial complex-valued coefficient array are allocated separately as illustrated in red and blue vertical strips. In such a way the real-valued curvelet coefficients can be obtained in the translated spatial locations.

In this schematic in Fig 17, each layer of concentric square represents an individual subband J and the each of the coefficient array denotes the orientations l of that subband. Each coefficient in the array designates the value of the curvelet coefficient at location (k_1, k_2). The curvelet transform implemented in the CurveLab toolbox [15] gives a considerable redundancy because the number of curvelet coefficients is much larger than the number of pixels. The redundancy is introduced by multiple coefficients in different orientations at same translation location (k_1, k_2). As shown in Figure 18, the curvelet coefficient at the same location in different arrays of one sector (North or East) are associated with the curvelet waveforms cantered at the same location but with respective coefficient magnitudes. Although it depends on the characteristics of a given image, in most cases, there is only one coefficient, or possibly two, having significant magnitude.

To well exploit the advantages provided by the highly-anisotropic curvelet and perform the texture analysis of the prostate pathological images based on the multiresolution method, we

consider the maximum curvelet coefficient at the translated (k_1, k_2), maximized with respect to all orientations, for texture analysis purpose in the sense of sparsity.
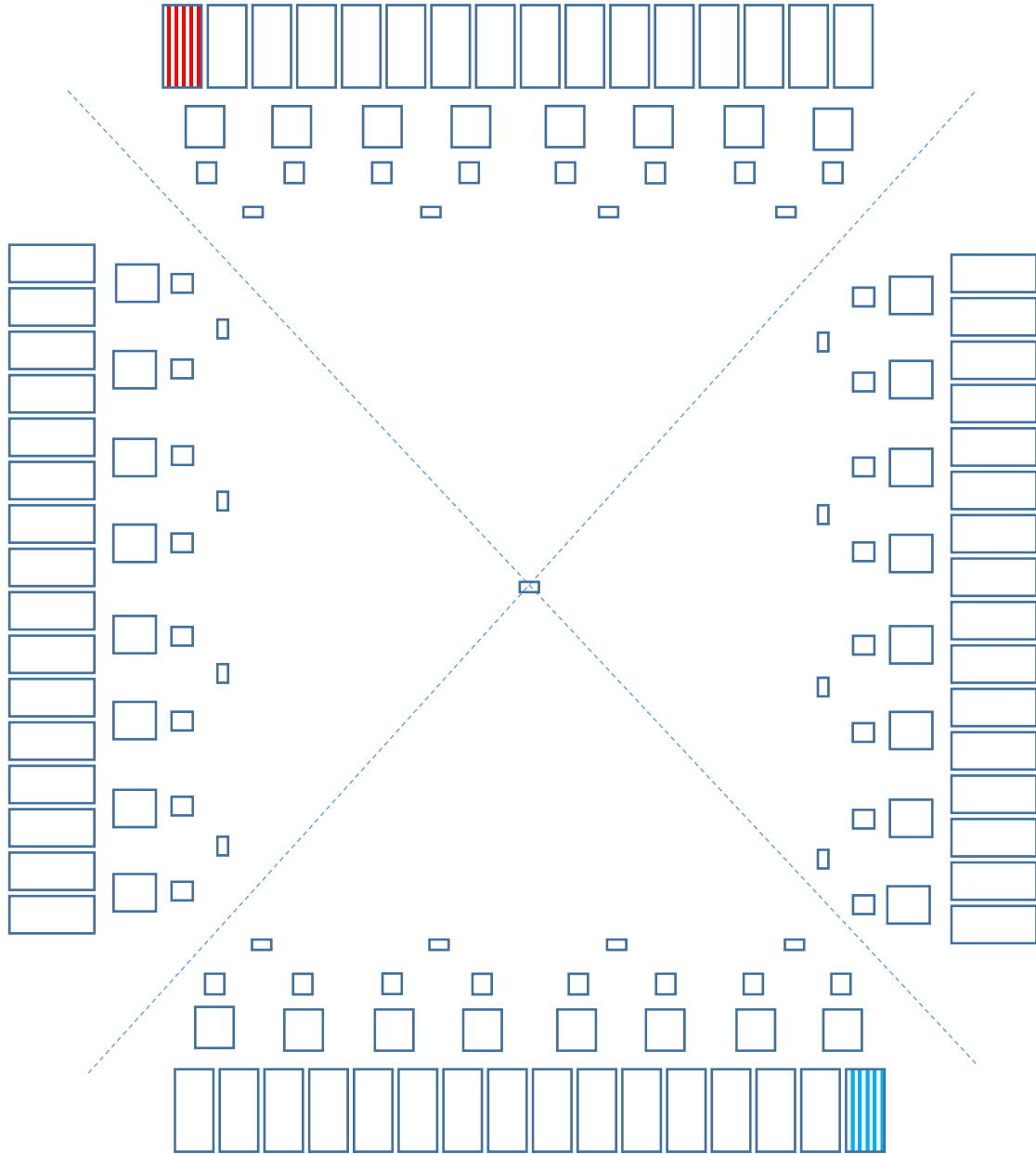
Figure 16. The schematic of curvelet coefficient arrays of a 256×256 pixel size image with 5 subband decomposition.
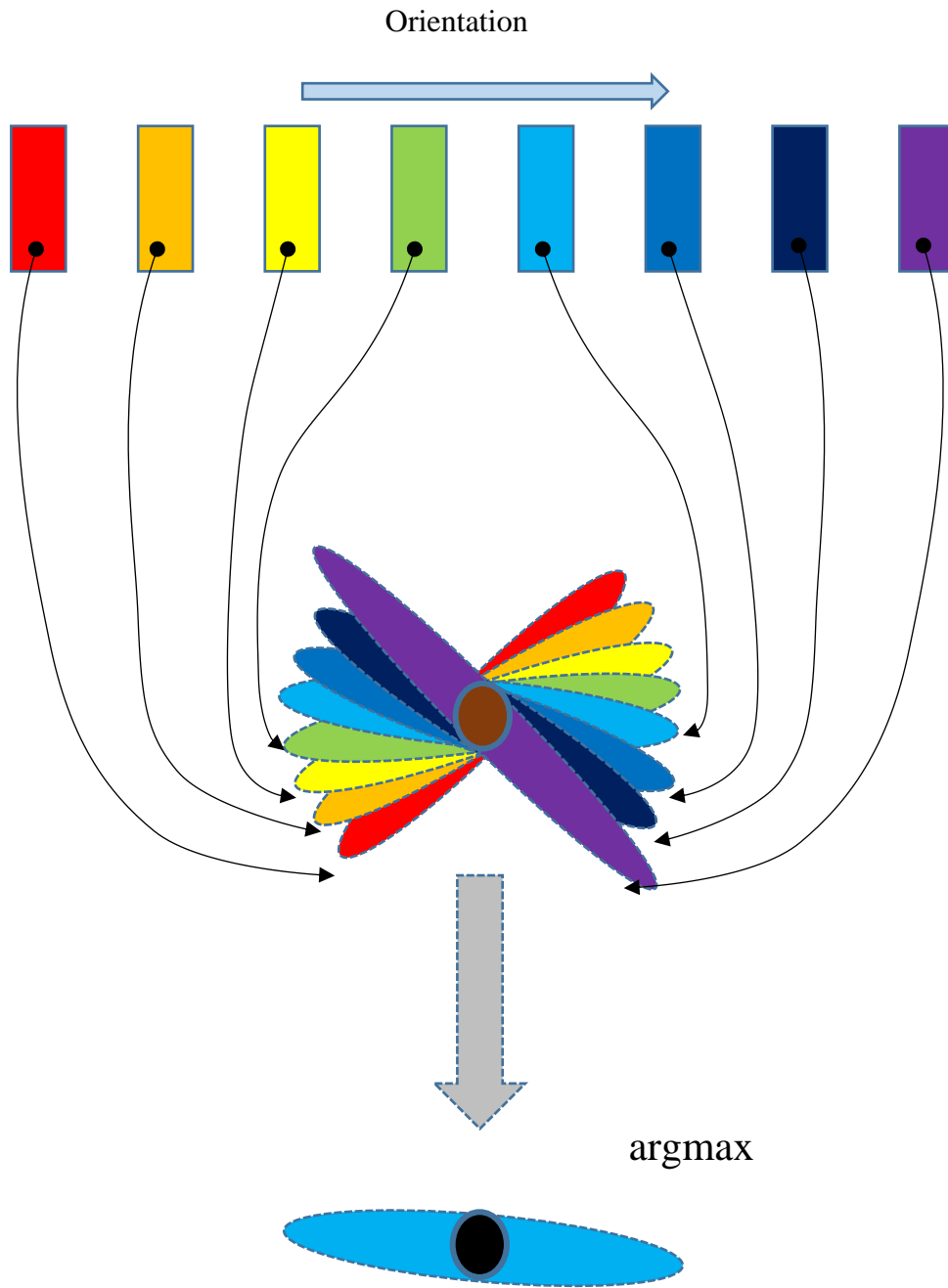
Figure 17. Illustration of the significant curvelet coefficient extracted from the coefficients of all arrays in one sector.
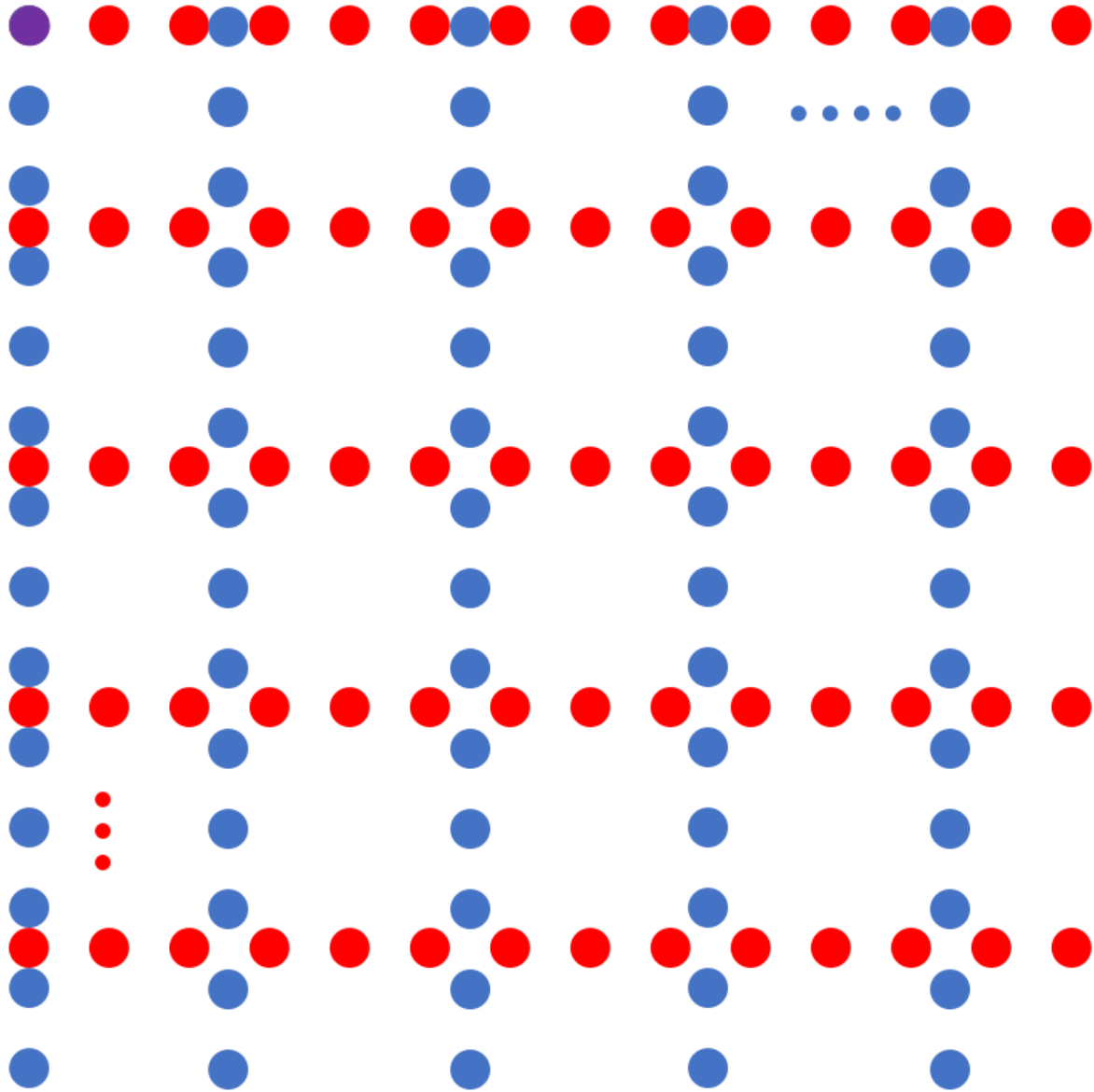
Figure 18. Illustration of the centers of curvelet supports or curvelet coefficient locations in coefficient arrays of scale $j = -1$ (subband $J = 5$) for an image of 256×256. Blue dots denote the locations in the arrays of the North sector while the red dots denote the locations in the arrays of the East sector.

Due to the parabolic scaling rule, curvelet coefficient at (k_1, k_2) or the translation k_1 and k_2 in the coefficient arrays in the North sector and in the East sector are swapped. When the inverse discrete Fourier transform is applied to the wrapped frequency components, the obtained curvelet coefficients in arrays of the North sector will be on the Cartesian grid $I_N = \left(2^j k_1, \quad 2^{j/2} k_2\right)$, $(j < 0)$, while that in the Easter sector will be on $I_E = \left(2^{j/2} k_1, \quad 2^j k_2\right)$. The curvelet coefficients in the arrays of the North sector will be dense along the vertical axis of the image and coarse along the horizontal axis, while the coefficients in the arrays of the East sector will be contrary.

Taking into consideration of the different resolution grids, the computation of the maximum curvelet coefficients at a location (k_1, k_2) will take two steps. First, we collect a set of coefficients at the same location closest to $\mathbf{k}$ from all arrays (covering an orientation range of $\pi/2$) in the North sector and designate it as $c_j^N(k, \theta)$. Similarly, collect another set of are the curvelet coefficients at the point nearest to $\mathbf{k}$ from all arrays in the East sector, where the resolutions of the translation grids are swapped, and designate it as set $c_j^S(k, \theta)$. The maximum curvelet coefficients at (k_1, k_2), for a given scale j is then given by

$$c_j^*(k) = \max_\theta \left| c_j(k, \theta) \right| := \max_\theta \{ c_j(k, \theta) \big| c_j(k, \theta) \in c_j^N(k, \theta) \bigcup c_j^E(k, \theta) \}. \qquad (3.1)$$

where $\mathbf{k} = (k_1, k_2)$ refers to the integer grids of the original image, and the maximization is taken with respect to all θ. Figure 19 shows the centers of curvelet supports where the blue dots represent the set of c_j^N and the red dots denote the set of c_j^E. Table 1 lists such a union set of centers of curvelets at which curvelet coefficients are assigned for scale $j = -1$ or subband 5 for a 256×256 image. As the curvelet coefficients of significant strengths provide the apparent curve segment

information in all scales and covering the whole image, the total number of the maximum curvelet coefficients in a 5 scale decomposition of an image of 256 × 256 pixels is only 10,297 and the redundancy now is reduced to about 0.3142 as compared to the number of pixels of the original image.

TABLE 1. Locations of curvelet coefficients in an array (scale $j = -1$, image size 256×256). Blue values represent the centers of the curvelet supports in arrays of the North sector while the red values denote the centers in arrays of the East sector

1, 1	1, 3	1, 5	1, 6.8182	1, 7	1, 9	1, 11	1, 12.6364	1, 13	1, 15	1, 17	1, 18.4546	1, 19	...	1, 251	1, 251.1818	1, 253	1, 255
3, 1			3, 6.8182				3, 12.6364				3, 18.4546				3, 251.1818		
5, 1			5, 6.8182				5, 12.6364				5, 18.4546				5, 251.1818		
6.8182, 1	6.8182, 3	6.8182, 5		6.8182, 7	6.8182, 9	6.8182, 11		6.8182, 13	6.8182, 15	6.8182, 17		6.8182, 19	...	6.8182, 251		6.8182, 253	6.8182, 255
7, 1			7, 6.8182				7, 12.6364				7, 18.4546		...		7, 251.1818		
9, 1			9, 6.8182				9, 12.6364				9, 18.4546		...		9, 251.1818		
11, 1			11, 6.8182				11, 12.6364				11, 18.4546		...		11, 251.1818		
12.6364, 1	12.6364, 3	12.6364, 5		12.6364, 7	12.6364, 9	12.6364, 11		12.6364, 13	12.6364, 15	12.6364, 17		12.6364, 19	...	12.6364, 251		12.6364, 253	12.6364, 255
...			...				...				...		...				
251.1818, 1	251.1818, 3	251.1818, 5		251.1818, 7	251.1818, 9	251.1818, 11		251.1818, 13	251.1818, 15	251.1818, 17		251.1818, 19	...	251.1818, 251		251.1818, 253	251.1818, 255
253, 1			253, 6.8182				253, 12.6364				253, 18.4546		...		253, 251.1818		
255, 1			255, 6.8182				255, 12.6364				255, 18.4546				255, 251.1818		

3.2.2 Maximum Curvelet Coefficient and Distributions

The discrete curvelet transform software generates curvelet coefficients of all orientations at a given location k, many of which are not significant that leads to an unnecessary high redundancy in contrast to the original objective of providing an efficient sparse representation. Based on the hypothesis that a given image function $f(x_1, x_2)$ may have a number of curved singularities at different locations but, otherwise, are smooth away from these singularities. At a given point (k_1,

k_2), only a unique curved singularity of a pertinent orientation exists, thus there will be only one curvelet coefficient there referring to that orientation. Instead of considering the use of a threshold to de-noise it [16], we propose to consider the maximum curvelet coefficient defined by

$$\text{Maximum curvelet coefficient} \quad c_j(k) = \max_\theta \left| c_j(k, \theta) \right|.$$

As given in eq. (3.1) in section 3.2.1, and develop our method of analysis and texture feature extraction.

Let us examine the curvelet transform of an example image shown in Fig. 19(a) and its maximum curvelet coefficients of three top scales in the decomposition (subbands 5, 4 and 3). The reconstruction from these three-scale maximum curvelet coefficients is shown in Fig. 19(b) displaying significant edge information around the constituent structures in the image. When it is added with the contributions from the scaling component and the subband 2 component, the result is shown in Fig. 19(c). This is the reconstruction of the image by using the maximum curvelet coefficients to obtain an approximation to the original image where the appearance of significant edge curves, regarded as textures, will help us to develop a effective new method for texture classification.
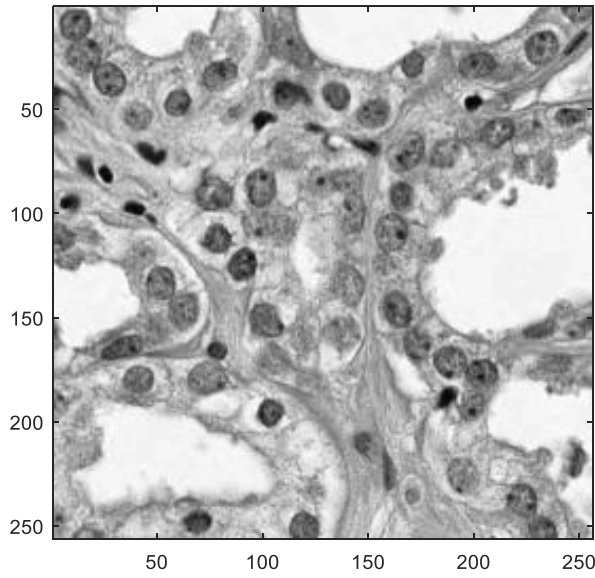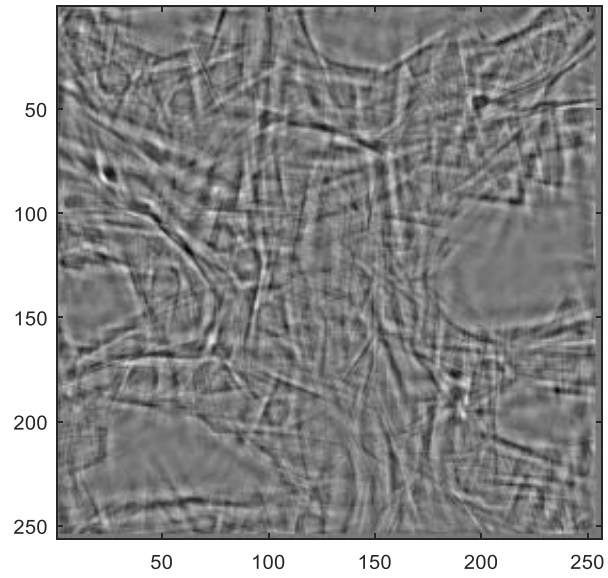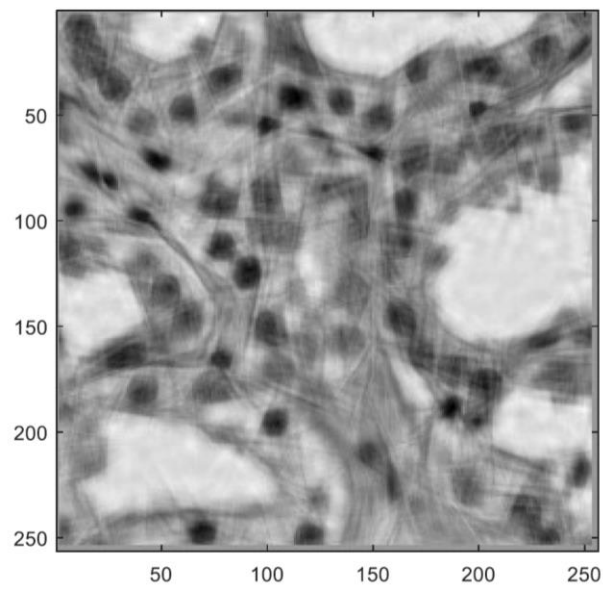
Figure 19. (a) an example tissue image, (b) reconstruction from maximum curvelet coefficients of three top scales,

and (c) reconstruction including contributions from scaling component and lower frequency component

The histograms of the maximum curvelet coefficients at each scale appear to have two-mode distributions in positive and negative coefficients, is almost symmetrical with respect to zero magnitude, all with strong curve strength. Each mode has sharp peak and is not symmetric with respect to its peak. The high peak of the original full coefficient distributions shown in Fig. 15, which delineates the slow change in image background is now vanished. Figure 20 shows the histograms of maximum curvelet coefficients of a P3S3 and a P4S4 image patch at scale from 3 to 5. They represent the physical boundaries of various nuclei and lumen more properly and provide, in reconstruction, a clear view of texture patterns displayed by these boundary curve segments. For small objects, both positive edges and corresponding negative edges always exist. Thus, the bi-modal histograms are almost symmetric. Let us reverse the sign of the negative coefficient magnitude and merge the negative mode to the positive side so as to give a single mode distribution of maximum curvelet coefficient magnitude which will still characterize the individual mode, yet give a simpler picture without having to calculate both modes separately.
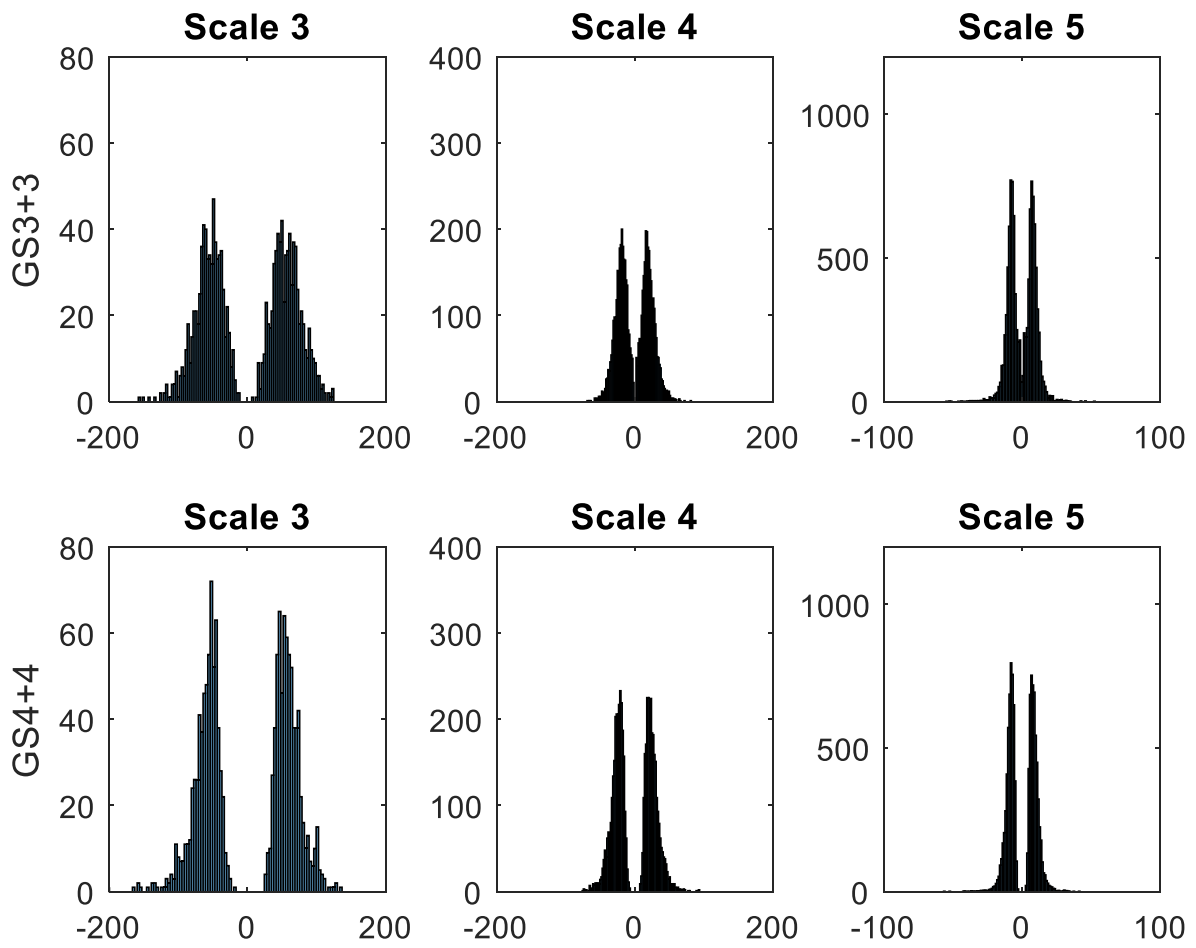
Figure 20. Histograms of significant curvelet coefficients

Figure 21 shows the histograms of the maximum curvelet coefficients of vertical, horizontal, diagonal and $\frac{3}{8}\pi$ orientations of a P3S3 and a P4S4 image patch at scale 4. From the histograms of the maximum coefficients of individual orientation, the distributions appear to have a similar skewness characteristic for Gleason grade 3+3 compared to that of grade 4+4. Figure 22 shows the histograms of the maximum curvelet coefficients of all orientations of a P3S3 and a P4S4 image patch at scale 3, 4 and 5. In the histograms of Gleason grade 3+3, there are more weak curvelet coefficients as compared to that in the histograms of Gleason grade 4+4. The close to zero magnitude coefficients represent areas of the well differentiated architecture that usually appear to be benign whereas cancerous areas that appear otherwise. In the histograms of Gleason grade 4+4 where the fusion of glands, increase number of cells, and obvious nucleoli with dark appearances in a cancerous region all contribute to give stronger curvelet magnitude. In this study, we focus on the texture features extracted from the maximum curvelet for classification of critical Gleason patterns.
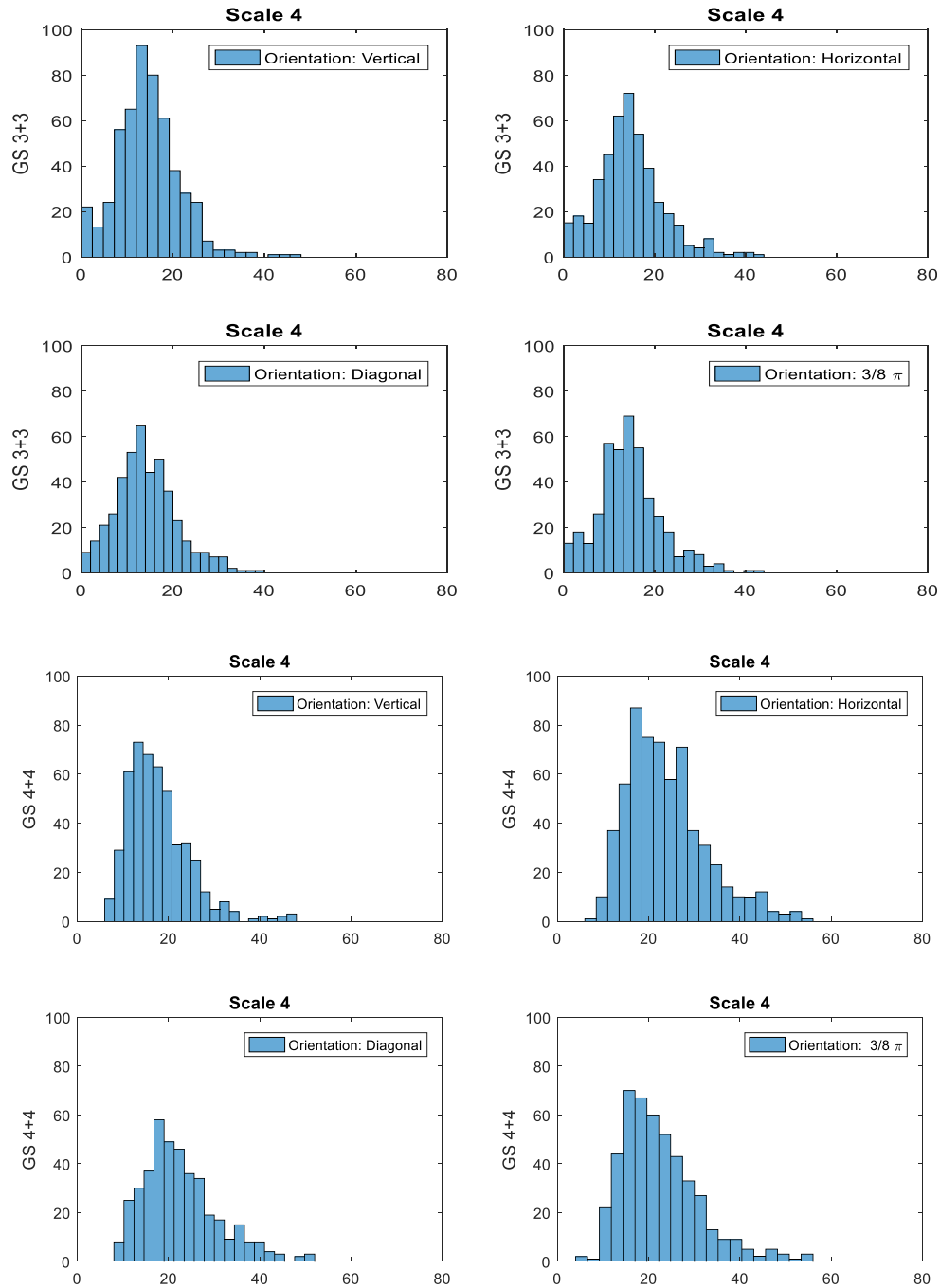
Figure 21. Histograms of maximum curvelet coefficients of three orientations in four orientations of a P3S3 and a

P4S4 image patches at scale 4 (with positive and negative coefficients pooled together)
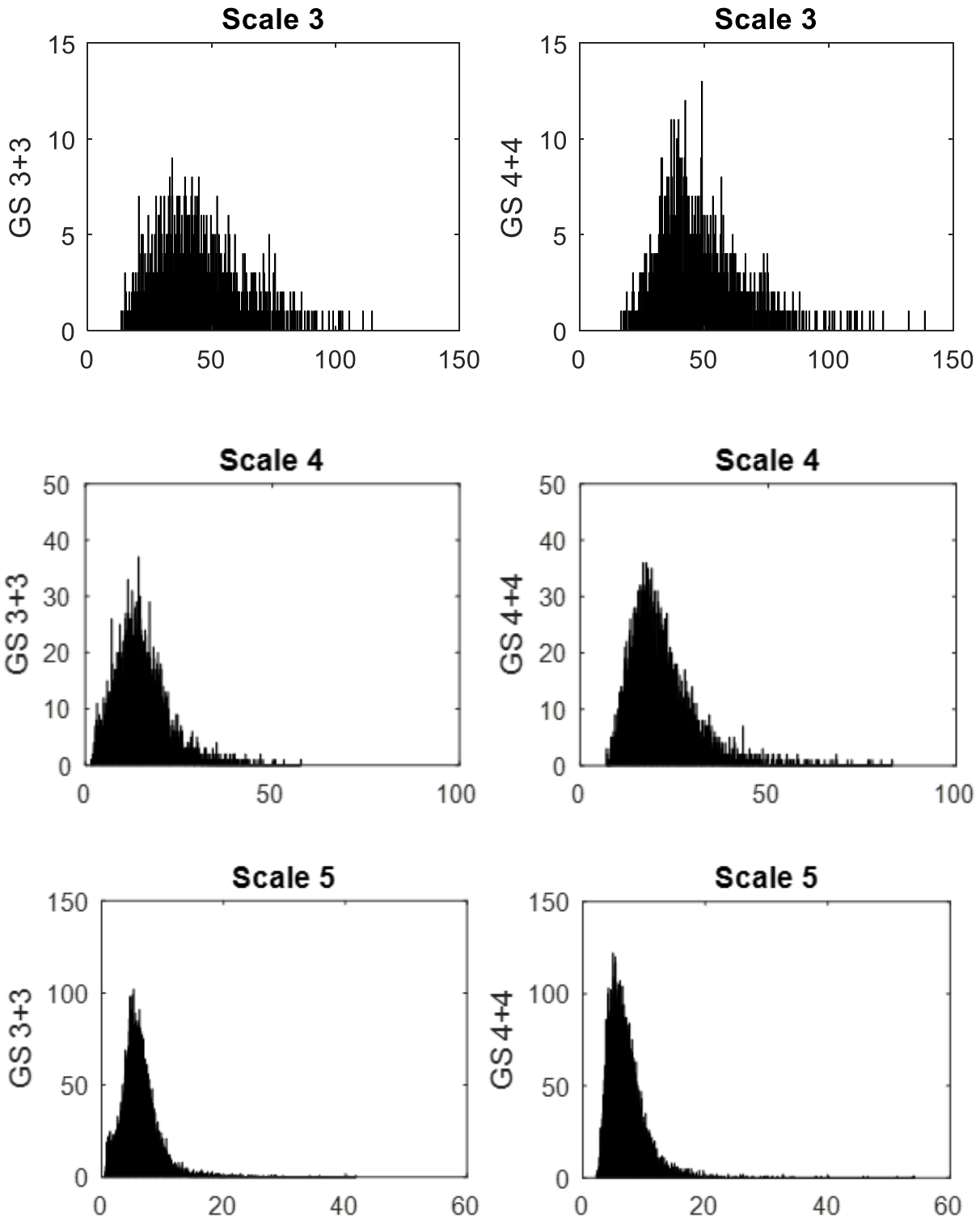
Figure 22. Histograms of maximum curvelet coefficients of a P3S3 and a P4S4 image patch at scale from 3 to 5

(pooling together coefficients in all orientations as well as of positive and negative magnitude)

The texture characteristics in an image can be used to estimate the associated architectures for prostate phenotype analysis. The maximum curvelet coefficients in 3 scales (subbands 5 to 3) of two sample images of primary Gleason score P3S3 and P4S4 are plotted in the lower part of Figure 23 and 24. The strength of coefficients is indicated by the brightness at their corresponding locations (k_1, k_2) in reference to the original image coordinates. The plot of low scale (subband 3) appears to be coarse as compared to that of the finer scale (subband 4 and 5). The maximum curvelet coefficients capture not only explicit micro textural information but also implicit macro structured information as well.

The texture information displayed in the fine scale curvelet coefficients plot (subband 5) of pure grade 3 shows the regions of benign cells have similar texture appearances while it is lacking in the case of pure grade 4. The curvelet-based coefficeints in fine scale are capable of capturing the subtle information such as cell, gland and stroma where the latter two components are more observed in grade 3 tissues but are less observed in grade 4.

As listed in Tab 2 and 3, the homogeneous texture information demonstrated by the curvelet-based features of the tissue image reveals the observation that either the patch is benign everywhere or malignant all the way through. Since the deteriorations of prostate cancer are associated with the transition of Gleason patterns from the purely normal Gleason pattern to cancerous phenotypes, this information can be measured in terms of the curvelet-based statistical texture features.
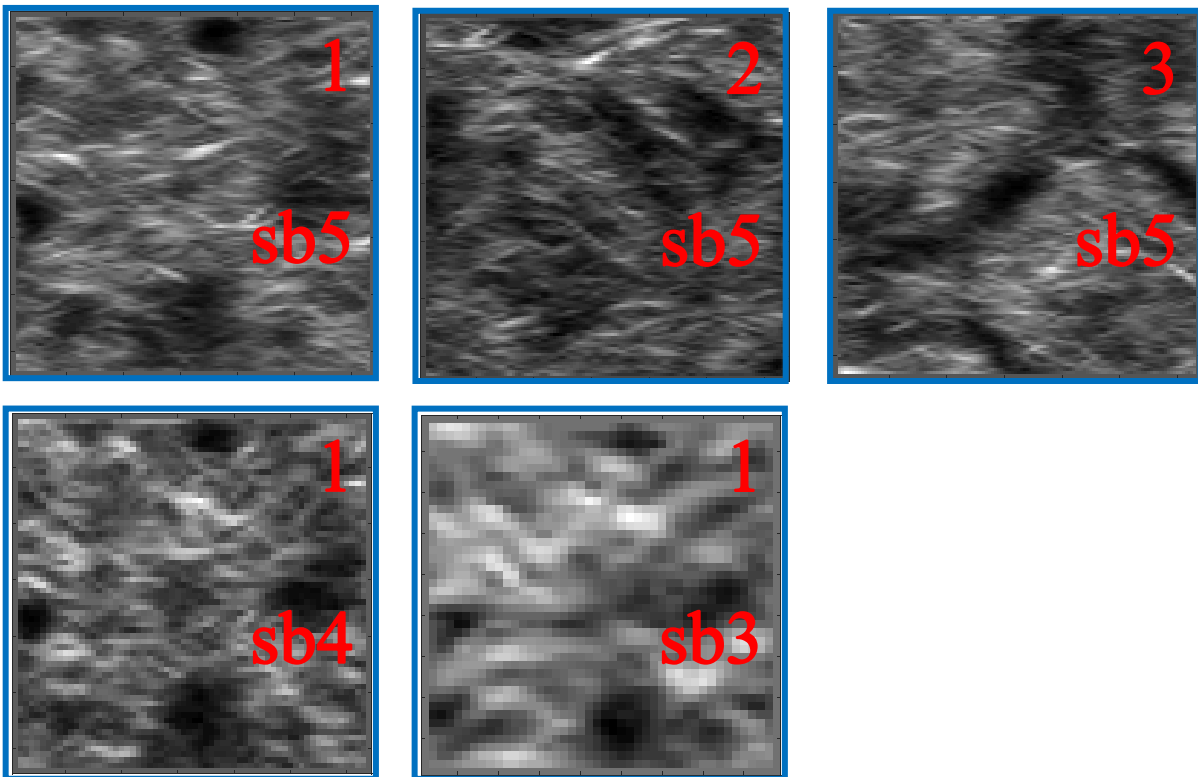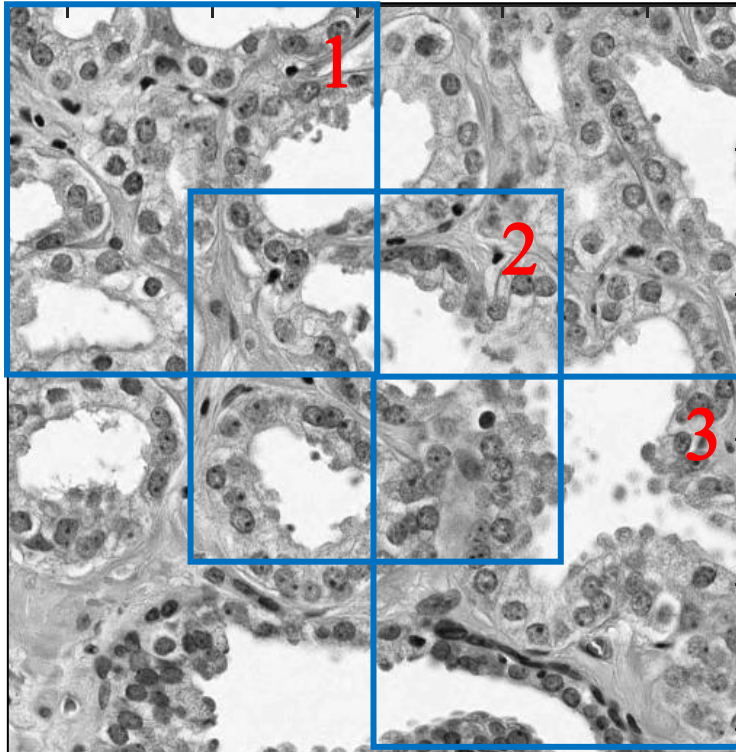
Figure 23. Maximum curvelet coefficients of three P3S3 image patches. The middle row shows the subband 5 and the bottom row shows different textures in subbands 4 and 3 of the first patch.
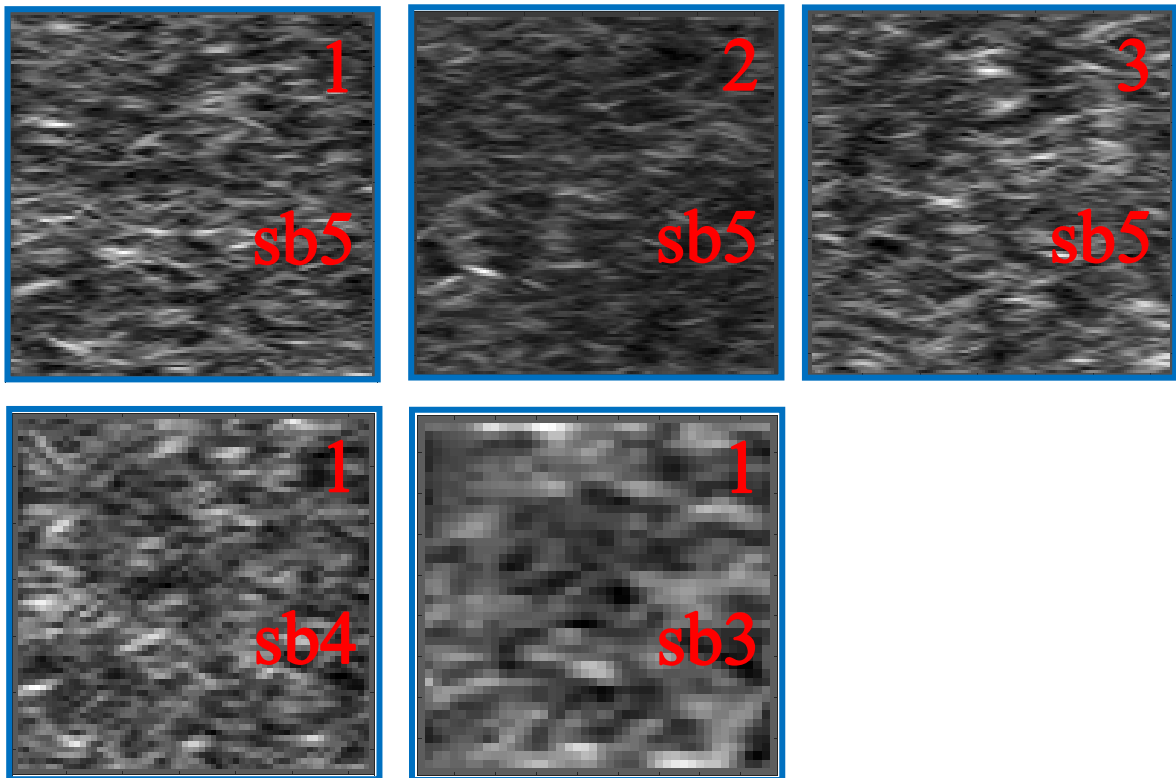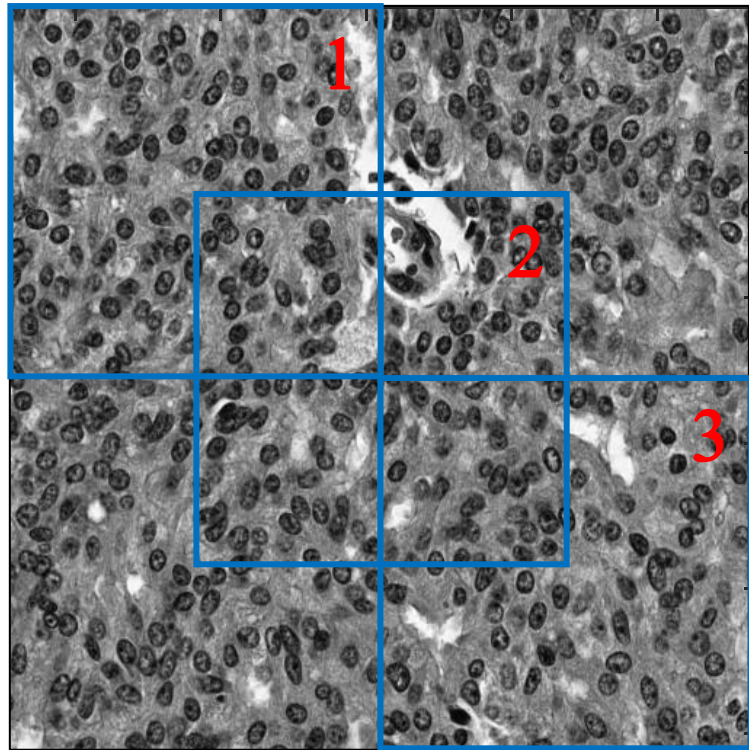
Figure 24. Maximum curvelet coefficients of three P4S4 image patches. The middle row shows the subband 5 and the bottom row shows different textures in subbands 4 and 3 of the first patch.
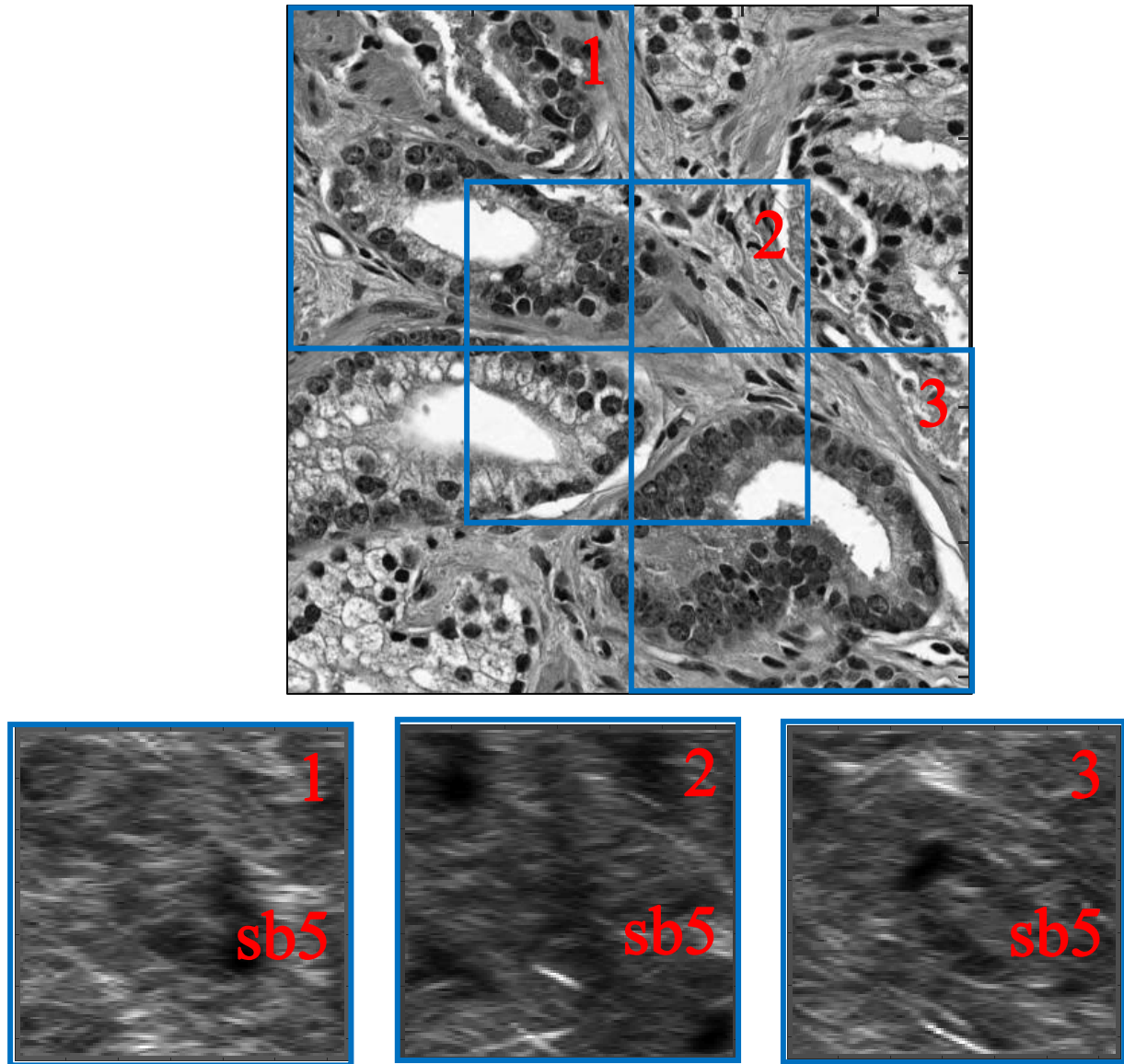
Figure 25. Maximum curvelet coefficients in subband 5 of three P3S4 image patches.
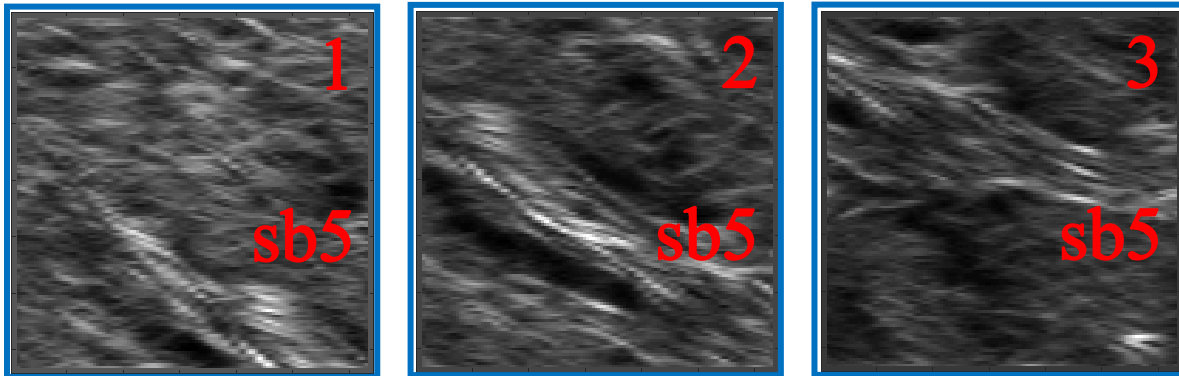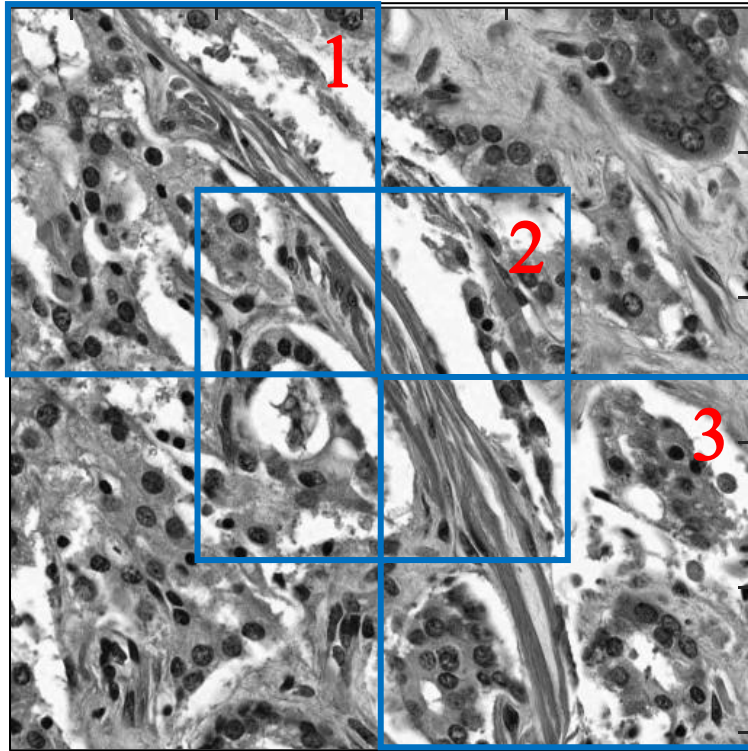
Figure 26. Maximum curvelet coefficients in subband 5 of three P4S3 image patches.

TABLE 2. Statistical features skewness in two scales of the image patches in Fig 23

	Skewness Subband 4	Skewness Subband 5
1	21.5533	13794.7315
2	21.7791	12935.7899
3	22.5651	14062.0205

TABLE 3. Statistical features skewness in two scales of the image patches in Fig 24

	Skewness Subband 4	Skewness Subband 5
1	25.8292	20451.6832
2	26.2017	20411.7193
3	25.9143	21473.8369

TABLE 4. Statistical features skewness in two scales of the image patches in Fig 25

	Skewness Subband 4	Skewness Subband 5
1	23.6037	53229.5336
2	24.4041	84339.5389
3	24.8680	78004.0870

TABLE 5. Statistical features skewness in two scales of the image patches in Fig 26

	Skewness Subband 4	Skewness Subband 5
1	25.0517	83272.7355
2	23.4383	142729.7029
3	23.2833	126941.7661

Azar et al. [109] proposed a statistical proximity sampling method to extract features from the sub-glandular structures in prostate tissue images which can represent the relative distribution of tissue components near lumen regions while preserving spatial quantitative information. It can be used as a basis for the analysis of different areas within an image. They obtained high classification rates for tubular formation in both healthy and cancerous tissue images which is considered as the most important information for the classification of prostate tissue and thus the method has demonstrated the efficacy in extracting discriminative features from the image patch. That is, the region-based statistical features extracted from the glandular structure comprising of lumen, epithelium, nuclei, stroma and associated components can provide effective tissue architecture and descriptions for the analysis of Gleason grading applied to the tissue prostate images. The implicit representation of region-based texture is robust and highly descriptive against the complexity of biological tissue and variations in staining procedures in different dataset.

The maximum curvelet coefficients of two sample images of intermediate Gleason score P3S4 and P4S3 are illustrated in Figure 25 and 26. The texture information displayed in the fine scale shows that these two in-between classes demonstrate different degrees of benign texture and cancerous appearance, which are heterogenous themselves, as compared to that of the primary grades. The comparison of the curvelet-based texture feature skewness of the fine scales listed in Table 4 and Table 5 indicates that it may be used to effectively discriminate the degrees of malignancy implied in the abundant textures. With such features extracted from the intermediate grades and the features obtained from the primary grades, we design a new classifier to achieve high accuracy as planned.

3.3 CURVELET-BASED TEXTURE FEATURE EXTRACTION

In the previous section, we have explored the texture characteristics exhibited in the multiscale curvelet coefficients of prostate cancer tissue images which contain abundant edge information of cells, nuclei and glandular structures and their arrangements. The curvelet magnitude carries much more information than pixel intensity, it denotes the strength of an oriented edge segment. It is a function of position variable and orientation variable. Its distribution is bi-modal and almost symmetric. At each scale, we may pool positive coefficient and negative coefficients together to give a single mode distribution for each orientation. We may also pool coefficients in all orientations together to assess a rotation-invariant distribution. Furthermore, we consider the maximum curvelet coefficient magnitude at each location, maximized over all orientations to achieve a sparse representation. These provide us a new basis for characterizing image textures for classification.

3.3.1 Texture Feature Extraction from Maximum Curvelet Coefficients

Based on the normalized histograms or probability distributions $P_J(c_J^*(k_1, k_2)) = P_J(k_1, k_2)$ of maximum curvelet coefficients $c_J^*(k_1, k_2)$ at each scale (subband J) with all orientations as well as positive and negative magnitudes pooled together, their first order statistics are used to extract the statistical texture features defined in the following:

- Variance:

Variance is a measure of how spread out the distribution of curvelet coefficient magnitude. A large variance value indicates that the coefficient magnitude in the image patch of $N \times N$ size are spread out widely.

$$\sigma_J^2 = \frac{1}{N^2 - 1} \sum_{k_1}^{N} \sum_{k_2}^{N} (c_J^*(k_1, k_2) - \mu)^2 P_J\left[c_J^*(k_1, k_2)\right]$$

$$\mu_J = \frac{1}{N^2} \sum_{k_1}^{N} \sum_{k_2}^{N} P_J\left[c_J^*(k_1, k_2)\right]$$

- Energy:

Energy is a measure of homogeneity of curvelet coefficient in the image patch.

$$Ener_J = \sum_{k_1}^{N} \sum_{k_2}^{N} P_J^2[k_1, k_2]$$

- Entropy:

The measure of complexity of the image. Complex textures tend to have higher entropy values.

$$S_J = -\sum_{k_1}^{N} \sum_{k_2}^{N} P_J[k_1, k_2] \log P_J[k_1, k_2]$$

- Skewness:

Skewness is a measure of symmetry of the distribution. A distribution is symmetric if it is the same to the left and to the right of the mean. The skewness for a normal distribution is zero, and any symmetric distribution should have a skewness very close to zero. Negative skewness value indicates that the distribution is skewed toward left of the mean which means that the left tail is long relative to the right tail

52

while positive value indicates skewed toward right of the mean. For the distribution of the maximum coefficients, only skewness toward right is apparent.

$$\gamma_J = \frac{1}{N^2 \sigma_J^3} \sum_{k_1}^{N} \sum_{k_2}^{N} \left| P_J\left[k_1, k_2\right] - \mu \right|^3$$

- Kurtosis:

Kurtosis is a measure of whether the distribution is peaked or flat compared to a normal distribution. Distribution with high kurtosis has a distinct peak near the mean, decays rapidly, and has heavy tails. A flat top near the mean rather than a sharp peak indicates that the distribution has low kurtosis.

$$kurt_J = \frac{1}{N^2 \sigma_J^4} \sum_{k_1}^{N} \sum_{k_2}^{N} \left| P_J\left[k_1, k_2\right] - \mu \right|^4$$

Figure 27 and 28 show the histograms of the curvelet-based textural features extracted from patches of prostate pathological images of two primary Gleason score 3+3 and 4+4, and two intermediate Gleason score 3+4 and 4+3, respectively. These features in different scales will be selected in our design of a classifier for a reliable recognition of the four critical Gleason scores.
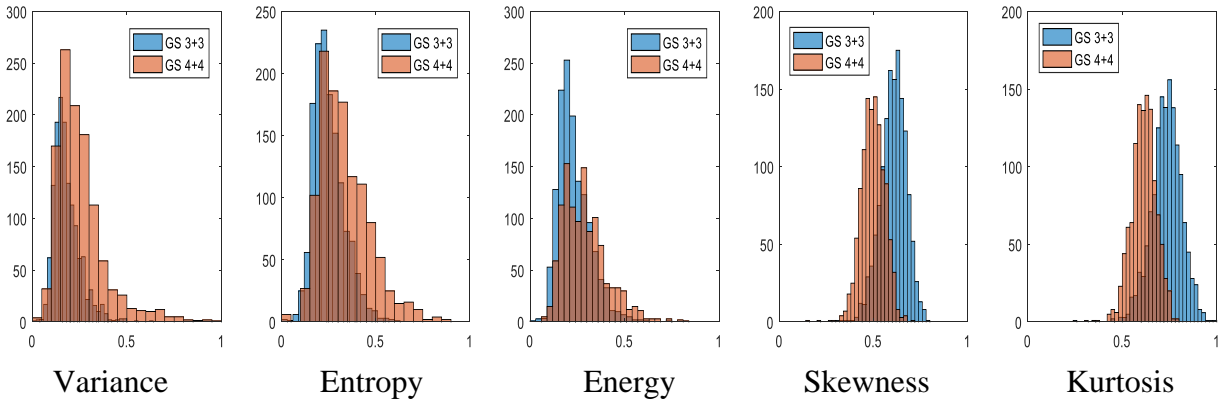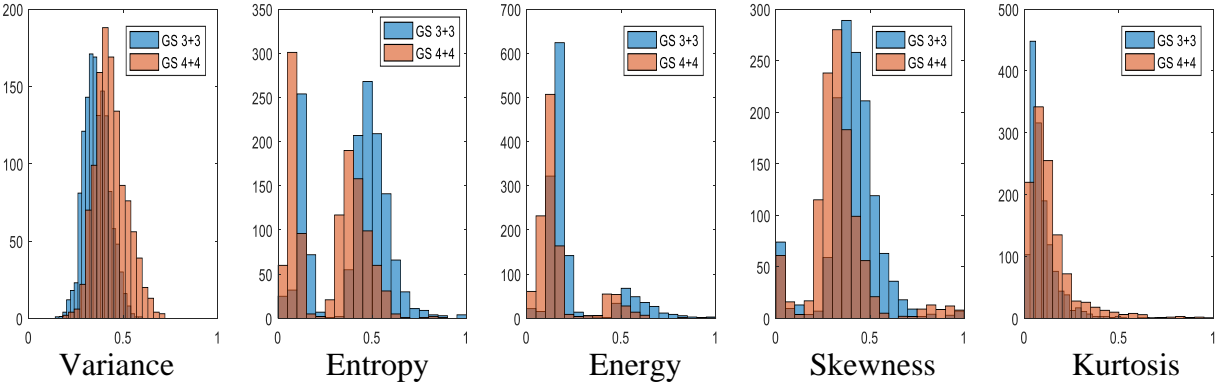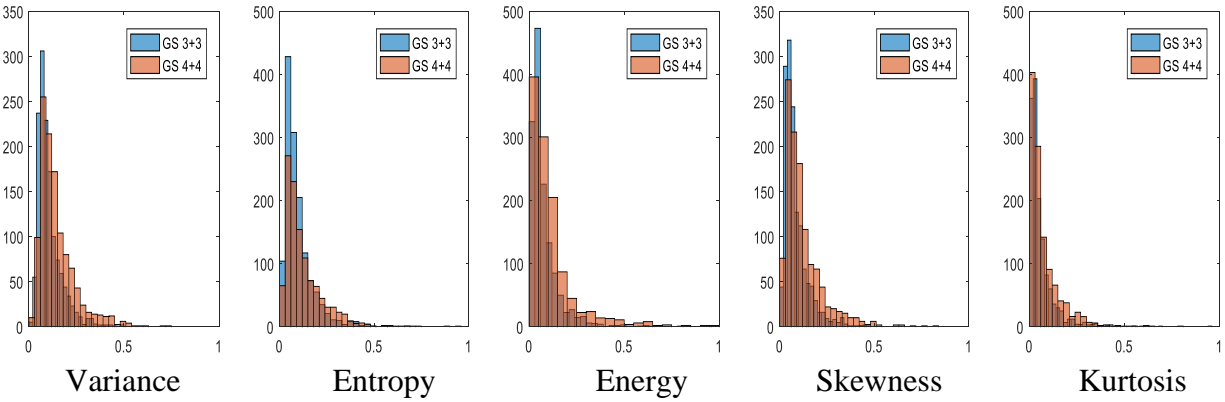
Figure 27. Histograms of statistical features extracted from maximum curvelet coefficients in (a) subband 3, (b) subband 4 and (c) subband 5 of two primary Gleason score 3+3 and 4+4.
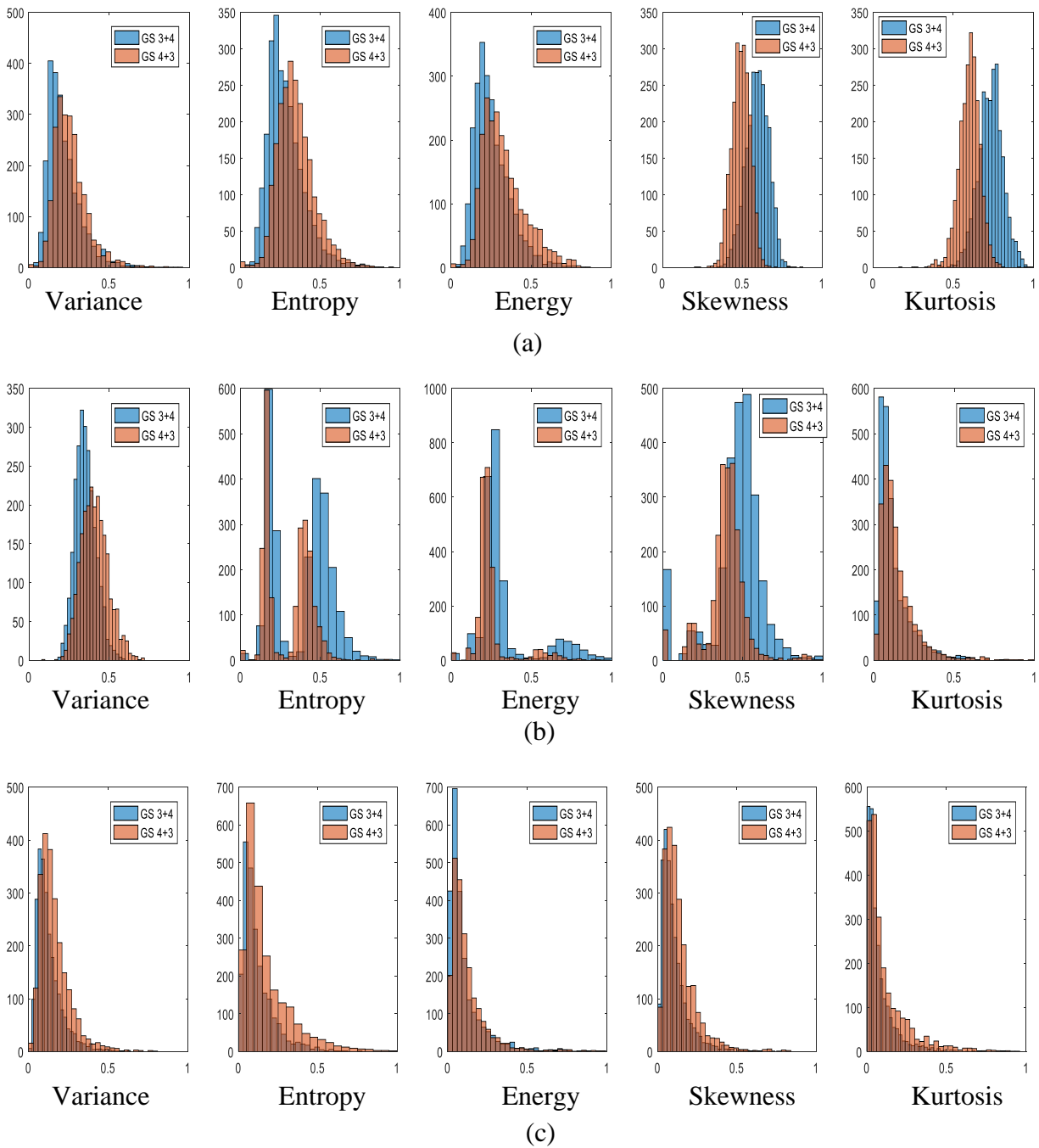
Figure 28. Histograms of statistical features extracted from maximum curvelet coefficients in (a) subband 3, (b) subband 4 and (c) subband 5 of two in-between Gleason score 3+4 and 4+3.

3.3.2   Dissimilarity Measure of Texture Features

The classical Kullback-Leibler divergence (KLD) can be used for the dissimilarity measure of feature x in two classes θ_i, $(i = 1, 2)$ which is given by [31, 68]

$$D(p(x;\theta_q)\|p(x;\theta_i)) = \int p(x;\theta_q)\log\frac{p(x;\theta_q)}{p(x;\theta_i)}dx. \qquad (3.6)$$

where $p(x; \theta_i)$ is the class-conditional probability density function of x. For the texture features x derived from maximum curvelet coefficients c_j^*, the density function estimates are generally non-Gaussian, the evaluation of KL-divergence would be more complicated. As a crude estimation by using Gaussian approximation, the divergence is simply given by

$$d_{1,2} = \frac{1}{2}\left(\frac{\sigma_2^2}{\sigma_1^2} + \frac{\sigma_1^2}{\sigma_2^2} - 2\right) + \frac{1}{2}(\mu_1 - \mu_2)^2\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right). \qquad (3.7)$$

where σ_i^2 and μ_i are variance and mean value of the feature x in each class. This divergence estimate is employed to provide a rough ranking and ordering of multiscale curvelet-based texture features in our investigation of feature selection based on their discrimination capability, that is to be presented in Chapter 4.

4.0 CLASSIFICATION OF FOUR CRITICAL GLEASON SCORES

This chapter presents new study on machine-aided classification of four critical Gleason patterns in prostate cancer tissue images by using texture features extracted from maximum curvelet coefficients of the images. As mentioned earlier, the reliable recognition of these patterns among Gleason score 6, 7 (3+4 or 4+3) and Gleason score 8 is of crucial importance that will affect the patient's treatment. A two-level classifier consisting of two Gaussian kernel support vector machines, each is incorporated with a voting mechanism by multiple windowed patches in a core image for class assignment in each level, has been developed. A set of Tissue MicroArray (TMA) images of four prominent Gleason scores (GS) 3 + 3, 3 + 4, 4 + 3 and 4 + 4 has been studied in machine learning and testing. The experimental result has achieved an average accuracy of 95.83% for 4 classes, an outstanding performance when compared with other published works.

4.1 STRATEGY IN CLASSIFIER DESIGN

When a urologist assigns each tissue image a Gleason grade number, primary grade and secondary grade respectively, the image will be firstly examined for its majority (> 50%) arrangement of glandular architecture (an area where the cancer is most aggressive) and then the secondary Gleason grade pattern where less extensive of growth (< 50%) is confirmed. In view of that, the

Gleason patterns are defined according to how well the glands are differentiated and the degree of cells are recruited by the surrounding neoplastic tissues. The P3S3 samples most often have individual discrete well-formed glands while the P4S4 samples typically have poorly-formed glands or predominantly lacking glands with a lesser component of well-formed glands [70-72]. Therefore, the perceptions of the homogeneity of the purity and impurity learned from prostate tissue images of pure grades, e.g. P3S3 and P4S4, can be used toward the interpretation of the heterogeneity displayed in images of the intermediate Gleason grades.

Based on the aforementioned characteristics of the growth of prostate cancer that involves structural and textural changes over time, we first investigate the discrimination between the primary Gleason score 3+3 and 4+4 by training a SVM classifier without the use of C parameter that controls the maximum penalty imposed on margin-violating observations. A SVM was trained with the pure grade 3 and 4 samples while the parameter C was not used. The resulting σ value and number of support vectors are listed in table 6 which demonstrates that the homogeneity displayed in the primary grades and captured by the curvelet-based texture features can be used to provide a principle margin to ascertain the recognition of grade 3 and grade 4. Although the training was successful finished with a proper selected σ value, however, the excessive number of support vectors will not contribute to a desirable generalization capability while the small sigma value would also indicate the over-fitting of the data points. This leads to the search and selection of hyperparameters including sigma and C values for a stable model with maximum gap tolerance. The SVM trained with the primary grades can be incorporated with a second SVM classifier to handle the heterogeneity and integrated into a four-class tree classifier towards a good classification result.

TABLE 6. Results for the SVM classifier

Classifier	σ	C	SVs
GS 3+3 vs 4+4	0.8993	-	339

The strategy of developing a 2-level tree classifier mimic the human expert's decision process. The classifier consists of one Gaussian-kernel SVM with a voting mechanism by multiple samples of image patches from a core image will give three categories in each level, as shown in Fig. 29. The SVM in the first level of the tree classifier is designed to learn the texture homogeneities appeared in pure grade 3 (Gleason score 3+3) and in pure grade 4 (Gleason score 4+4), respectively, from training patch samples. If the training is 100% successful, the 25 patches in an image must have the same texture pattern for the particular Gleason grade. The classification results of these 25 patches must be unanimous, either grade 3 or grade 4, which implies that the image as a whole is of Gleason score 3+3 or 4+4. For a testing image, if the voting is not unanimous, some patches vote for grade 3 and some vote for grade 4, then the image will be assigned to the indecision category (including GS 3+4 and GS 4+3). That means, the 25 patch classification results provided by this SVM#1 are used as 25 votes to decide whether the image has an absolute majority of votes of either pattern 3 or of pattern 4, or else has a heterogeneous mixture. The absolute majority rule applied here requires 100% of the votes, i.e., 25/0 (all Gleason 3 patterns) for class P3S3, and 0/25 (all Gleason 4 patterns) for class P4S4. Once the classification outputs indicate that there is a discrepancy among the 25 patches, i.e., from 24/1 to 2/23, then these 25 patches will be directed to the second level of the tree classifier where a machine SVM#2, which is trained by GS $3 + 4$ and GS $4 + 3$ training samples, will make a decision. Among the 25 classification results generated by these patches underwent SVM#2, the voting majority ($\geq 14/11$) will assign the image to the

class 3 + 4, the voting majority ($\leq 11/14$) will assign it to the class 4 + 3, and in-and-between cases 13/12 and 12/13 are allocated to the Indecision category.

Our goal is aim at achieving a better separation of the critical Gleason score which has to be done by consolidating the classification of the primary grade first through the recognition of texture embedded therein. Since the texture pattern of GS 3+3 and GS 4+4 image will only be either pure grade 3 or 4, the first SVM in the layer one of the tree classifier is designed to constrict these texture patterns toward their groups as close as possible and thus the design of the Indecision group is crucial. If any of the impurity has been found in any one of the 25 image patches, the particular image section containing the 25 patches will be classified as the Indecision category and directed to the next level classifier where the in-between classes are discriminated according to different degree of mixture lies in them.

That is, although the training of the SVM could be finished successfully with the use of slack variables and margin parameter C, when one considers a novel image sample, the trained tree classifier can only assure the accuracy on the patterns it learned before and thus any one of the 25 patches that falls to the wrong side of the decision hyperplane should be treated strictly to ensure the correct classification. In a similar fashion, the second SVM of the tree classifier is also designed with a safety mechanism that strictly forestalls unrecognized patterns from going into the decision process. The Indecision category reinforces the separation hyperplane in a way to prevent misclassification of the critical Gleason grade a step further. The proposed tree classifier is very distinct from the standard multi-class support vector machines.

Gaussian kernels were used in both SVMs where the parameter σ was chosen with a Bayesian optimization procedure, in each kernel for both machines. Individual SVM has different number of support vectors as resulted from the training. The two kernel SVMs were successfully

trained, together with the successful majority decision rule, giving a trained tree classifier of 4

classes of critical Gleason scoring with no training error.
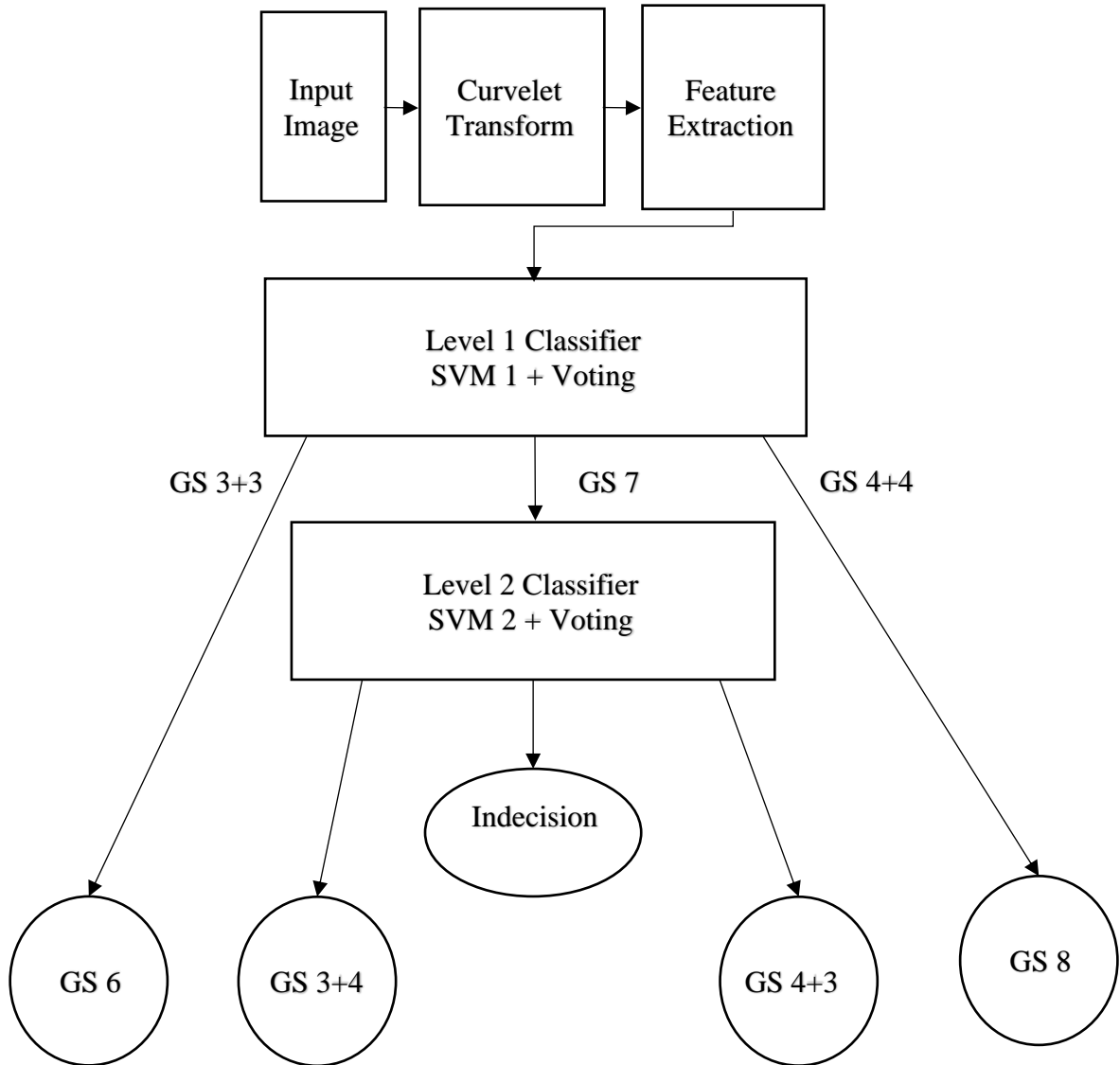
Figure 29. The proposed tree-structured classifier for critical Gleason grading.

4.2   METHODS

The curvelet-based methodology for statistical texture feature extraction from prostate pathological images is described below. Non-linear support vector machines are developed for classification of critical Gleason score 3+3, 3+4, 4+3 and 4+4.

4.2.1   Sampling Image Patches

A central portion of 768×768 pixels covering about 70% of the core was taken which was considered to be sufficient for enclosing the prostate cells and glandular structure information. Let a window of the size of 256×256 pixels be moved across the whole 768×768 array starting from the top-left area to the bottom-right each time shifted by one half of the width (128 pixels) columnwise or rowwise, thus a total of 25 patches are taken with partial overlapping. Each patch of 256×256 pixels undergoes the fast discrete curvelet transform [12, 13] with the use of the Curvelab Toolbox software [15] to give curvelet coefficients $c(j,l,k)$ in the decomposed subbands where the specific spatial frequencies are wrapped and the inverse Fourier transform is performed to provide the oriented edge information of the prostate cellular and ductile structures contained therein.

Figure 30 is shown for the image patches obtained with moving window method, taken from a GS 4+4 images. Their statistical measures including variance σ_J^2, energy Ener_J, entropy S_J, skewness γ_{1J}, and kurtosis Kurt_J in each subband for each patch are computed as textual features as we discussed in Chapter 3.
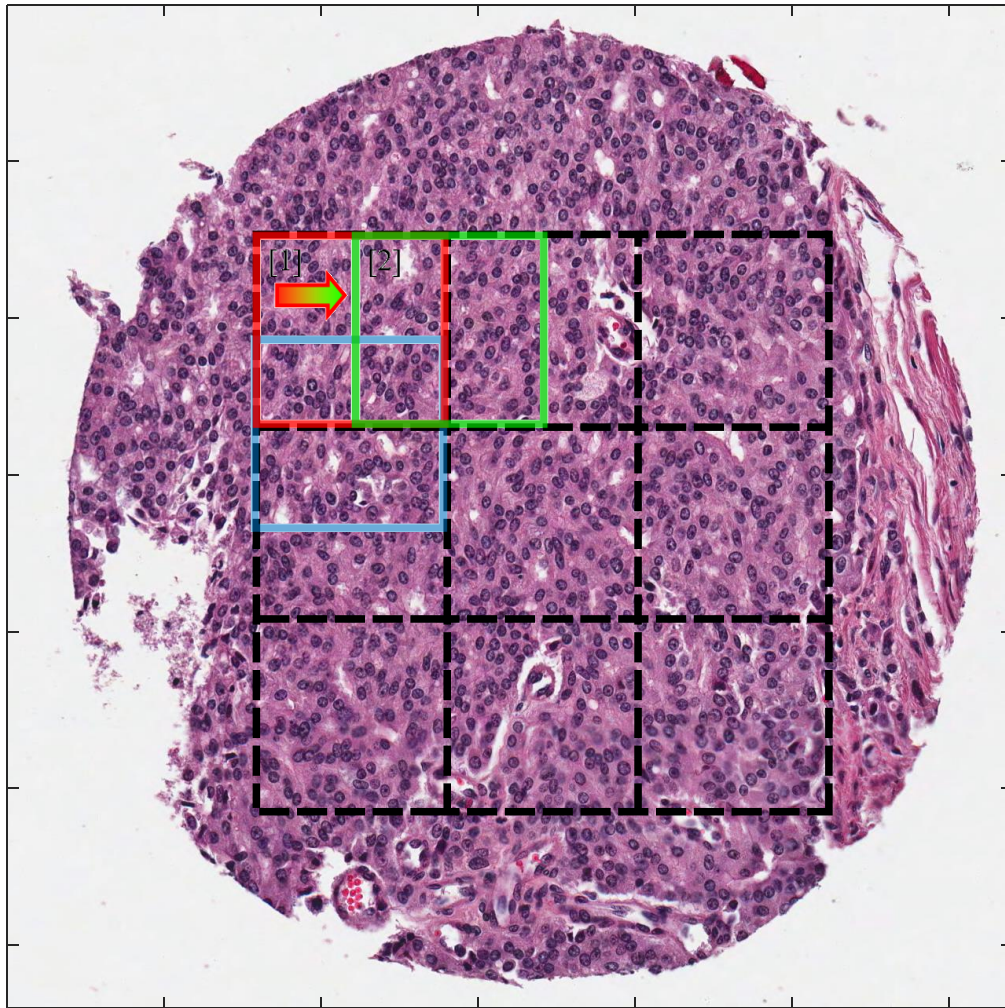
Figure 30. Moving window samples of (a) P3S4 and (b) P4S3 TMA images

4.2.2 Feature Selections

Each TMA image is about 2.8 megapixel color images and comprises a biopsy core of 0.6mm in diameter captured by a microscope at 20× magnification. We examined 768×768 pixels of the central part constituting about 75% area of the section which contains abundant prostate cellular and glandular information. It is subdivided into 25 smaller patches, each of 256×256 pixels with

partial overlapping column-wise or row-wise. In other words, a scanning window of this size samples different portions of the image 25 times to form 25 local observations. Each observation will be tested by the proposed tree classifier and the majority of the 25 testing results are used to decide which Gleason patterns the image belongs to.

Our previous work [19, 97-98] has showed that the use of statistical features obtained from the conventional curvelet-based textures and a proper designed tree classifier can achieve a preferable classification performance compared to current published research where the critical intermediate Gleason grades are less reported. A tree structured classifier consisting of three Gaussian-kernel support vector machines (SVM) was developed where the first SVM is initially to decide whether an input patch belongs to Grade 3 (GG3 denotes the inclusion of P3S3 and P3S4) or Grade 4 (GG4 denotes the inclusion of P4S3 and P4S4) and then to make majority decision of multiple patches in the image. The result of the first machine will direct the image to either the second SVM or the third SVM, where machine#2 is responsible to decide whether the image belongs to class P3S3 or class P3S4, and machine#3 to decide whether it belongs to class P4S3 or class P4S4; each machine likewise examines multiple patches and makes a majority decision.

The maximum coefficients obtained by applying the fast curvelet transform to the available prostate histological images with moving window method are pulled together in each of the 5 decomposition scales. Statistical measure including variance, energy, entropy skewness and kurtosis are calculated from these coefficients. According to the contemporary Gleason grading system [38, 70, 72], grade group 1 (Gleason score ≤ 6) will only have individual discrete well-formed glands while the grade group 2 (Gleason score $3 + 4 = 7$) and above grade groups start to deteriorate with poorly-formed, fused or cribriform component and finally lose the gland formation.

Cancer cells can particularly be distinguished from non-cancer cells by changes in the nucleolar structure. Alterations in nucleolar shape, size and amount are common features in quantitative nuclear morphometry that has been used to predict progression of different types of cancer [70]. The nucleolar features can be detected by the curvelets and quantized into statistical measures for efficient analysis. From well-differentiated to poor-differentiated adenocarcinomas, the mean number nucleoli increase significant and the main morphologic changes in the structure of nuclei are also characteristic features of cancer cells. In recent years, quantification of these cell features has been used to evaluate development of different cancers [74] and the studies have shown that the alterations of nuclear structure can be used to predict metastasis and progress in prostate cancer [75]. For prostate carcinomas evaluation, information extracted from nucleolar morphometry that measures nucleolar number, area, perimeter, and diameter can forecast progression as well as biochemical recurrence as these size and shape changes are frequent events in prostate cancer cells and are usually associated with a more aggressive cancer phenotype.

In tissue images, nucleoli in a benign prostate gland usually appear consistently dark or uniformly bright over its entire area without obvious nucleoli, while nuclei in a cancerous gland appear light blue and contain prominent nucleoli which appear as small dark spots [74]. Cancer cells can be particularly distinguished from non-cancer cells by looking at the changes in the nuclei structure. Alterations in nuclei shape and size and increase in the number of nucleoli are common features in quantitative nuclear morphometry that have been used to predict metastasis and forecast progression as well as biochemical recurrence in prostate cancer [74-76]. The nuclei features can be examined via the curvelet transform of the image and quantized by statistical measures of the curvelet coefficients where the strength reflect edges of different components and their spatial

arrangements give textural patterns. These can provide more effective classification capability as developed in our work.

The images contained in the database are prostate cancer histological images of Gleason Scores 6, 7 and 8, consisting of 4 classes: P3S3, P3S4, P4S3 and P4S4. The texture information provided by the tissue images could exhibit the homogeneous and heterogeneous characteristics implicitly or explicitly depending on which class the cancer is assigned to. This kind of information can be used towards the learning of the subtle difference of critical cancer patterns and thus for the discrimination of these categories. The machine vision tasks will need to be able to handle the decision flow similar to that of pathologists. This can be fulfilled through the recognition of multiple patches of subimages covering each image area. In order to process the purity and impurity exhibit in prostate tissue images, a tree structure classifier that consists two support vector machines (SVM) embedded with voting mechanism is developed.

We evaluate the KLD method of curvelet-based texture feature selection for the classification of critical prostate cancer histological images. Total 10 curvelet-based textural features from the fine scales 3 to 5 have been ranked and provided as two reference feature sets for the classification of primary Gleason score 3+3 versus 4+4 and 3+4 versus 4+3. Among the two sets, the same 7 out of the 10 features are selected according to the Kullback–Leibler divergence method with different ranking for the two categories while 3 extra features, the Kurtosis and Entropy of subband 5, and Variance of subband 4, are selected to provide the discriminant information for the second classification task which involves of certain degree of mixture of the primary patterns. The feature set for the classification of primary Gleason score 3+3 versus 4+4 is chosen as

$$\begin{bmatrix} S_4 \\ \gamma_4 \\ \mathrm{Ener}_4 \\ \gamma_5 \\ \mathrm{Ener}_5 \\ \mathrm{kurt}_4 \\ \sigma_3^2 \\ \mathrm{kurt}_3 \end{bmatrix},$$

while the feature set for the classification of primary Gleason score 3+4 versus 4+3 is chosen as

$$\begin{bmatrix} S_4 \\ \mathrm{Ener}_4 \\ \gamma_4 \\ \sigma_4^2 \\ \mathrm{kurt}_5 \\ S_5 \\ \gamma_5 \\ \mathrm{Ener}_5 \\ \mathrm{kurt}_3 \\ \sigma_3^2 \end{bmatrix}.$$

The methodology for an effective classifier design with ensemble approach that realized by weighted voting mechanism will be developed and evaluated. The experimental results to illustrate how feature suggested by the KLD method and effective classifier design that improve the classification performance will be discussed and reported in next chapter.

4.2.3 Gaussian Kernel Support Vector Machines

Suppose that a training set for the 2-class problem consists of N training samples of d-dimensional feature vectors $x_i \in \mathbf{R}^d$ and the class label be $t_i = +1$ for class 1 and -1 for class 2. The training samples are denoted by $\{(x_i, t_i)\}_{i=1}^N$. (The linearly non-separable case can be better addressed by using the SVM with a kernel function to map the input vectors from the original space into a high dimensional hidden space where an optimal separating hyperplane can be constructed.) If they are linearly separable, let a linear decision hyperplane be given by

$$g(x) = w^T x + b = 0 \qquad (4.1)$$

where w is the weight vector and b is the bias or threshold weight. We may fine two hyperplanes called carinal hyperplanes, one on each side of the separating hyperplane $g(x) = 0$ with a distance, any training sample lying on either one of the cardinal hyperplane is called a support vector denoted as x_{SV}. The margin area between the separating hyperplane and the cardinal hyperplane may be considered as a bound. The goal for training a SVM is to find the optimal decision hyperplane that gives the largest bounds. With appropriate scaling the bound can be normalized to 1 so that,

$$
\begin{aligned}
w^T x_i + b \geq 1 \quad & \text{for} \quad x_i \in class \quad 1, \quad t_i = +1 \\
w^T x_i + b \leq -1 \quad & \text{for} \quad x_i \in class \quad 2, \quad t_i = -1
\end{aligned}
\qquad (4.2)
$$

The margin distance can be maximized which is equivalent to minimize the $\frac{1}{2}\|w\|^2$. The training of SVM is to find the optimal w and b using the training sample set subject to the constraints in Eq. (4.2) i.e.,

$$t_i\left(w^T x_i + b\right) \geq 1 \qquad \text{for } i = 1,2,...,N \tag{4.3}$$

To allow some training sample points on the opposite side of its cardinal hyperplane, we may introduce a non-negative slack variables ξ_i then the objective function becomes

$$Q(w,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \tag{4.4}$$

which is to be minimized with respect to w where $C > 0$ is a regulation parameter and the separating hyperplane must satisfy

$$t_i\left(w^T x_i + b\right) \geq 1 - \xi_i, \qquad \xi_i \geq 0 \quad \text{for } i = 1,2,...,N \tag{4.5}$$

The term $-\sum_i \alpha_i\left[t_i\left(w^T x_i + b\right) - (1 - \xi_i)\right] - \sum_i \mu_i\xi_i$ must be added to the right side of objective function Q in eq. (4.4). The $\sum_{i=1}^{N}\xi_i$ is the upper bound of the number of the erroneous samples and a large C leads to a large penalty for such an error while a small C yields a high error rate on the training vectors [104]. This offers a trade-off between maximum margin and minimum error. The constrained optimization problem leads to the discriminant function expressed by

$$g(x) = \text{sign}\left(\sum_{i \in x_{SV}} \alpha_i t_i x_i^T x + b\right) \tag{4.6}$$

where x_{SV} are the support vectors. When the feature vectors are nonlinearly mapped by a function $\Phi(x): \mathbf{R}^d \rightarrow H$, an optimal decision hyperplane can be constructed in that space as

$$g(x) = \text{sign}\left(\sum_{i \in x_{SV}} \alpha_i t_i \Phi^T(x_i)\Phi(x) + b\right) \tag{4.7}$$

Let a kernel function K be defined as $K(x_i, x) = \Phi^T(x_i)\Phi(x)$ and the optimal hypersurface is now given by

$$g(x) = \text{sign}\left(\sum_{i \in x_{SV}} \alpha_i t_i K(x_i, x) + b\right) \tag{4.8}$$

If $\Phi(x_i)$ is a Gaussian function, then the Gaussian kernel is given by

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right) \tag{4.8}$$

The kernel parameter σ is the standard deviation which determines the width of the Gaussian distribution. The Statistics and Machine Learning Toolbox was used for training our SVM classifiers [108].

To achieve a good generalization capability and attain a low error rate on future data, the appropriate values of the cost parameter C and the kernel parameter σ, the hyperparameters used by the SVM have to be estimated. Hyperparameters cannot be directly identified and thus a train and cross-validation process is required to evaluated them recurrently. In general, a grid search in (C, σ) space to obtain the optimal object function value is performed for the model selection [107, 105] and the performance of the hyperparameter optimization is measured with cross-validation on the held-out set. The hyperparameters that achieve the best accuracy in the train-validation cycle are finalized as the model parameters. The Bayesian optimization [106-108] are used in the search of the hyperparameters in this work.

4.3   EXPERIMENTS AND RESULTS

The TMA prostate histological images used to train the classifier are from pathology databases prepared by the Urology/Pathology Laboratory of the Johns Hopkins Medical Institute. Among these files (TMA #471, #681 and #682), 32 available images from 16 cases (patients) of GS 3 + 3 and GS 4 + 4, respectively were used in the training of Gleason patterns of Grade 3 and Grade 4 which provided 64 core images in total where 1,600 (= 25×64) training patch samples were generated by the aforesaid local scanning window approach. The 1,600 training feature vectors were used to train the SVM 1 in the first level of the tree classifier for classification of the patch images observations and then casting votes of 25 sub-images in each case for determining pattern 3 or pattern 4 of the core images. The support vector machine 2 in the second level of the tree classifier is trained with 32 images from 16 cases of GS 3+4 and GS 4+3 data sets, respectively, to provide also 1,600 training feature vectors for pattern discrimination of patches and then use majority balloting to differentiate pattern P3S4 from P4S3. For the first and second SVM's in the classifier, feature sets for P3S3 vs P4S4, and for P3S4 vs P4S3, respectively, are listed in Table 7, they were selected according to Kullback–Leibler divergence criteria [68]. The digits J in entries of Table 7 represent the subband J where the features were extracted from. The higher order moments, skewness and kurtosis, were selected as texture features for their discriminating capability, as discussed in previous sections.

There are 8 features used in the first SVM 1 including variance and kurtosis from subband 3, energy, entropy, skewness and kurtosis from subband 4, energy and skewness from subband 5. The 10 texture features used in the second SVM 2 include the same 7 features used in the first

SVM but with three additional features from fine subbands, variance from subband 4, entropy and

kurtosis from subband 5, to provide enhancement of discrimination for the in-between classes.

TABLE 7. Statistical features in different subbands (n) selected by the KL-Divergence method for the tree classifier

Classifier \ Feature	Variance	Energy	Entropy	Skewness	Kurtosis
GS 6 vs GS 8	3	4, 5	4	4, 5	3, 4
GS 3+4 vs GS 4+3	3, 4	4, 5	4, 5	4, 5	3, 5

4.3.1   Kernel Selection

The model selection for the RBF kernel used in the two support vector machines in the tree

classifier was carried out by the Bayesian optimization method [105]. The input feature data were

normalized into [0, 1] and the search ranges for RFB kernel σ and regularization parameter C were

from 10^{-1} to 10 and from 1 to 10^4, respectively. For each pair of the hyperparameters, a leave-

one(image)-out (LOO) cross-validation test was performed. Then we selected the value of the

hyperparameters that achieved the best recognition rate as the model parameter values. Table 8

lists the results of the selected values for σ and C. From the table, the regularization parameter C

selected for SVM 2 in the second level is larger than that of SVM 1 in the first level. The same

situation was obtained with the resulting number of support vectors (SV) in the two SVMs, the

second one is much large than the first one. The objective function value and the CV error versus

the number of times of function evaluations are shown in Figure 31-32. The object function of the

first SVM converged faster than that of the second SVM as shown in Fig. 31, and the first SVM

also reached a lower cross-validation error rate as compared to the second SVM. Number of

support vectors, values of alpha, and threshold weights of the two SVMs are listed in Table 13 and

Table 14 in Appendix B, respectively.

TABLE 8. Validation results of the tree classifier

Classifier	σ	C	Error	SVs
SVM#1	1.5545	1	1.12%	72
SVM#2	2.7974	6.1212	4.42%	171
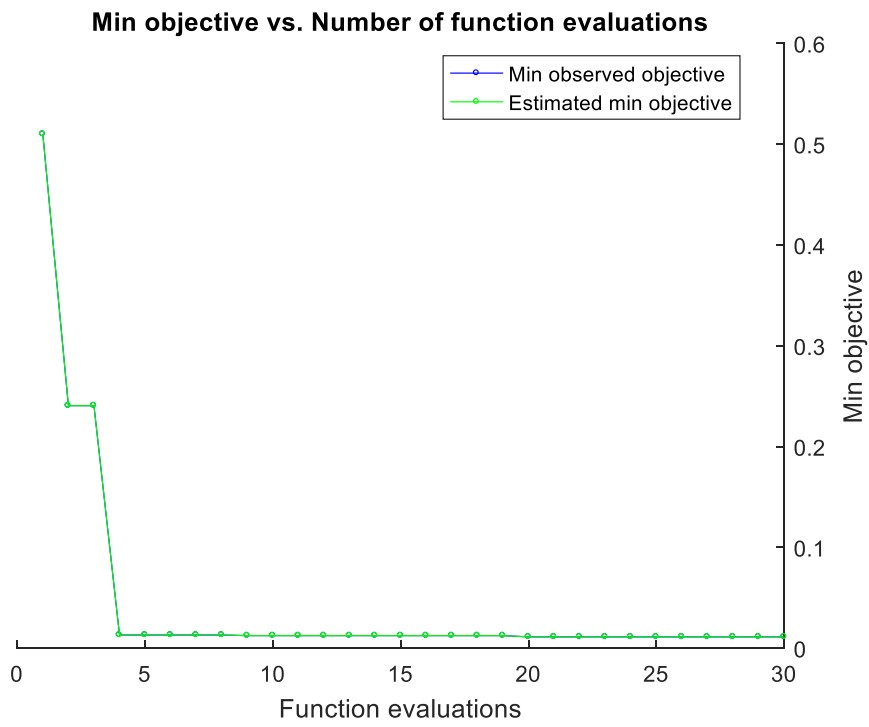
Figure 31. The cross-validation error versus evaluation times of the first SVM.
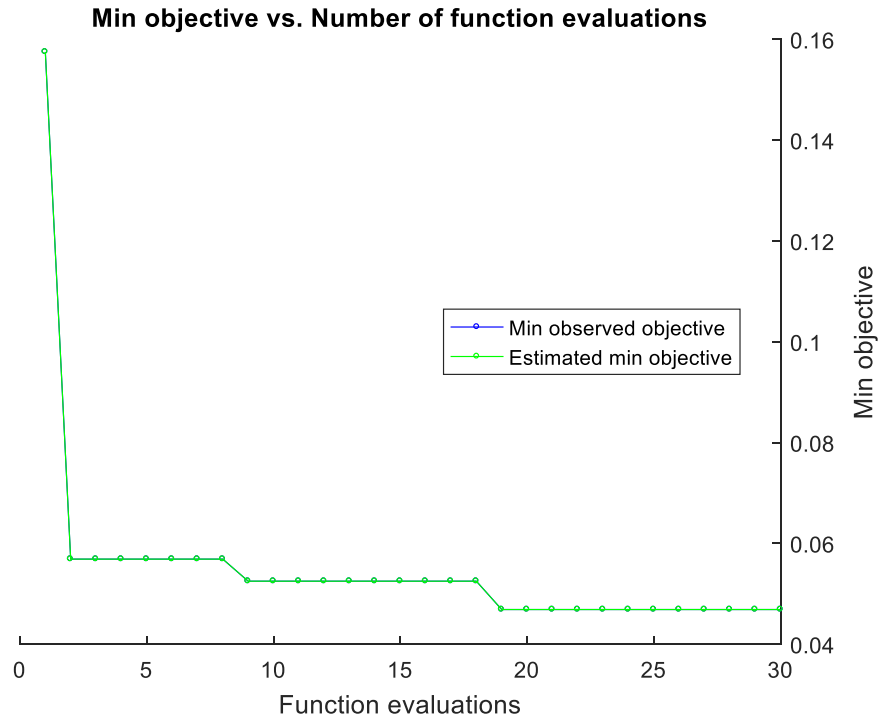
Figure 32. The cross-validation error versus evaluation times of the second SVM.

4.3.2 Training and Testing Results of the Tree Classifier

Using the Gaussian support vector SVM 1, the training of 32+32 core images of GS 3+3 and of GS 4+4 was 100% successful as shown in Table 9. To evaluate the generalization capability for this level classifier, firstly, we conducted the 10% Jackknife cross-validation tests for 100 realizations and the resulting statistics are shown in Table 10 with an average accuracy of 98.88% for primary Gleason grade 3 and 4. We also tested the SVM 1 with 32 core images of Gleason 3+4 and 32 core images of Gleason score 4+3 which would be used as training set for SVM 2 in the level 2. The results were given in Table 10 and listed in Table 11. All the images of Gleason score 3+4 and 4+3 were correctly classified to class 3 (Gleason score 7) at level 1 of the tree classifier.

74

TABLE 9. Training results of SVM 1 (based on image patches)

Train Label	GS 3+3	GS 4+4
GS 3+3	32×25	0
GS 4+4	0	32×25

TABLE 10. Validation test of SVM 1 (based on image patches)

Test Label		GS 3+3	GS 4+4	Error	Overall Accuracy
Class 1	GS 3+3	3165	32	1.09%	98.88%
Class 2	GS 4+4	37	3163	1.16%	
Average		98.90%	98.84%	1.12%	

Level 1 classifier SVM 1 + patch voting

Input with training samples of GS 3+3 and 4+4				Input with samples of GS 3+4 and 4+3		
Test Label		Class 1 GS 3+3	Class 2 GS 4+4	Class 3		Error
				GS 3+4	GS 4+3	
Class 1	GS 3+3	32	0	0	0	0%
Class 2	GS 4+4	0	32	0	0	0%
Class 3	GS 3+4	0	0	32	0	0%
	GS 4+3	0	0	0	32	0%

TABLE 11. Testing result of trained (SVM 1 + voting) with a set of samples GS 3+4 and 4+3

No \ Grade	GS 3+4		GS 4+3	
	3	4	3	4
1	14	11	10	15
2	17	8	11	14
3	13	12	12	13
4	13	12	8	17
5	14	11	12	13
6	22	3	12	13
7	15	10	4	21
8	14	11	10	15
9	16	9	7	18
10	13	12	6	19
11	13	12	9	16
12	13	12	11	14
13	18	7	8	17
14	17	8	12	13
15	16	9	10	15
16	18	7	10	15
17	13	12	3	22
18	14	11	4	21
19	24	1	12	13
20	21	4	9	16
21	21	4	5	20
22	16	9	9	16
23	15	10	9	16
24	13	12	7	18
25	21	4	10	15
26	13	12	9	16
27	14	11	11	14
28	17	8	11	14
29	14	11	6	19
30	13	12	4	21
31	18	7	8	17
32	14	11	12	13

Secondly, we conducted the training of SVM 2 in the second level with 1600 training image patches of each class, which was also 100% successful and then the 10% Jackknife cross-validation tests for 100 realizations. The results are given in Table 12 and Table 13, respectively. The resulting statistics shown in Table 13, gave an average accuracy of 96.95% for Gleason score 3+4 and 4+3.

TABLE 12. Training results of SVM 2 (based on image patches)

Train \ Label	GS 3+4	GS 4+3
GS 3+4	32×25	0
GS 4+3	0	32×25

TABLE 13. Validation results of SVM 2

Test \ Label	GS 3+4	GS 4+3	Indecision	
			GS 3+4	GS 4+3
GS 3+4	3102		98	
GS 4+3		3103		97
Average	96.94%	96.97%	3.06%	3.03%

The capability of the trained tree classifier was evaluated by a test set of TMA images which includes 24 images in 12 cases of each of the GS 3+3, 3+4, 4+3 and 4+4 in Table 14, gives the test result of the level 1 classifier (SVM 1 + voting) for the four classes. The samples assigned to class 3 with indecision labels at level 1 were directed to SVM 2 in level 2 for further classification. The overall testing results of the whole tree classifier are given in Table 15 with an average accuracy 95.83% for the four classes. The use of features generated from maximum curvelet coefficients which integrate both nuclei and glandular edge information into the texture

features appears to improve the classification accuracy of the four critical classes from 96.88%

validation accuracy reported in our earlier work [104] to the current descent (validation accuracy)

of 97.91%.

TABLE 14. Testing Result of the SVM 1

Grade No	GS 3+3		GS 3+4		GS 4+3		GS 4+4	
	3	4	3	4	3	4	3	4
1	25	0	12	13	10	15	0	25
2	25	0	16	9	11	14	0	25
3	25	0	13	12	12	13	0	25
4	25	0	16	9	12	13	0	25
5	25	0	14	11	13	12	0	25
6	25	0	13	12	10	15	0	25
7	25	0	14	11	11	14	0	25
8	25	0	14	11	10	15	0	25
9	25	0	14	11	11	14	0	25
10	25	0	13	12	12	13	0	25
11	25	0	15	10	12	13	0	25
12	25	0	13	12	9	16	0	25
13	25	0	13	12	9	16	0	25
14	25	0	14	11	10	15	0	25
15	25	0	13	12	12	13	0	25
16	25	0	14	11	10	15	0	25
17	25	0	14	11	9	16	0	25
18	25	0	13	12	12	13	0	25
19	25	0	15	10	11	14	0	25
20	25	0	15	10	11	14	0	25
21	25	0	13	12	13	12	1	24
22	25	0	14	11	12	13	0	25
23	25	0	14	11	12	13	0	25
24	25	0	12	13	10	15	0	25
	100%		91.67%		91.67%		95.83%	
Accuracy	95.83%				93.75%			

TABLE 15. Testing Result of the Tree Classifier

Test Label		Gleason Score 6 GS 3+3	Gleason Score 8 GS 4+4	Gleason Score 7 GS 3+4	Gleason Score 7 GS 4+3	Indecision	Overall
GS6	GS 3+3	24					24
GS8	GS 4+4		23		1		24
GS7	GS 3+4			22		2	24
GS7	GS 4+3				23	1	24
Accuracy		100%	95.83%	91.67%	95.83%		96
Accuracy						Average Accuracy	95.83%

4.4   DISCUSSION

The new tree classifier having two Gaussian kernel support vector machines with voting by multiple patches in a prostate tissue image using curvelet-based texture features of the images has been developed with success. The notion of a moving window to sample multiple subimages for use in majority voting is mimic to a pathologist's decision process for Gleason scoring attempting to roughly model the concept of the primary and secondary Gleason grading of a core prostate cancer histological image. The unanimous votes in level classifier during the training is to ascertain the feature characteristics of pure grade3 and pure grade 4. As demonstrated in our experiments, it is a very rewarding success of our effort.

The classification performance attained and the limited number of support vectors of SVM 1 in the tree classifier confirmed the effectiveness of selected features extracted from the high order statistical moments of maximum curvelet coefficients.

The classification of all 24 test images of Gaussian score 3+3 reached 100% accuracy. No test image of other scores was under-rated to GS 3+3. This is our objective.

79

The developed tree classifier has been demonstrated with the testing result giving an overall average classification accuracy of 95.81% for 4 classes. The first level SVM 1 with voting mechanism was designed on the basis of homogeneous textures of all patches in each training image of Gleason score 3+3 and similarly, 2 different but also homogeneous textures of Gleason score 4+4, respectively. The second level SVM 2 with majority voting instead of unanimous voting utilizes additional features in fine scales to enhance the discrimination between Gleason scores 3+4 and 4+3. The 95.83% and 91.67% recognition accuracy of GS 4+3 and GS 3+4 test images, respectively, involved 1 and 2 images, respectively being assigned to indecision category results were too close to the mid-point, but they stayed in Gleason score 7. One of the 24 test images of GS 4+4 was misclassified to GS 4+3. It was assigned to class 3 at the level 1 and subsequently at level 2 was classified into GS 4+3 where the classification was designed to only differentiate between GS 3+4 and GS 4+3. If we relaxed the requirement of unanimous (0/25) voting of GS 4+4 in level 1 classifier, e.g., using the threshold of (2/23), this error will be avoided.

The use of SVM classifier has been proved to be superior to the conventional KNN or Bayes classifier of Gleason grading problems [35, 65]. Compared to the reported results listed in Table 16, our new method has shown a remarkable performance. In [35], a sophisticated two-stage classifier which is composed of a multi-class Bayes classifier and binary SVM classifiers trained with multi-resolution quaternion wavelet transform (QWT) features and LBP (Local Binary Pattern) features, produced a comparable accuracy 98.83% in classifying Grade 3 and Grade 4, while ours is 98.88%. and our classifier is free from any error of misclassifying GS 4+4 to GS 3+3 and is capable of providing the classification of intermediate GS 3+4 and 4+3 classes which is a challenging task for histopathology image analysis [99, 109].

TABLE 16. Comparison of cross-validation of different Approaches for G3 vs G4 and 4 critical Gleason grades

Method	Dataset	Grade 3 vs Grade 4	GS 7 (3+4) vs GS 7 (4+3)
Cardinal multiridgelet transform (CMRT) texture [56]	16 grade 3 and 26 grade 4 images (20×)	93.75%	
Lumen- and nuclei-based method [109]	134 grade 3 and 70 grade 4 images (20×)	87.30%	
Shape features of the lumen and gland inner boundary [110]	16 grade 3 and 11 grade 4 images (40×)	95.19%	
Texton forests [34]	25 grade 3 and 50 grade 4 images (10×)	94.00%	
Network cycle features [98]	25 grade 3 and 50 grade 4 images (10×)	91.56%	
Gland segmentation and structural features [58]	28 grade 3 regions and 20 grade 4 regions (20×)	85.60%	
Quaternion wavelet transform (QWT), quaternion ratios, and modified LBP [35]	30 grade 3, 30 grade 4 and 11 grade 5 images	**98.83%**	

Texture features from combining diffusion coefficient and T2-weighted MRI images [99]	34 GS 3+3 vs 159 GS ≥7
	 114 GS 3+4 vs 26 GS 4+3
	 159 GS ≥7 includes:
	   114 GS 3+4,
	   26 GS 4+3,
	   19 GS ≥8
		93.00%	**92.00%**

Our two-level classifier using maximum curvelet coefficient-based texture and features	32 GS 3+3, 32 GS 3+4, 32 GS 4+3, and 32 GS 4+4 images (20×)	**98.88%**	**95.58%**

5.0   CONCLUSIONS AND SUGGESTIONS

One of the major efforts in this dissertation research is the utilization of the significant curvelet coefficient at each location which is the maximum coefficient with respect to all orientations at each point for sparse and realistic representation of images. As the coefficient at a point refers to a unique orientation, it denotes a piece of short edge segment there which may be a texton related to the image structure, for example, nuclei boundary or glandular structure in a tissue image. Extraction of texture features based on the significant curvelet coefficients is direction sensitive and robust to noise, and thus is more effective on representing texture characteristics.

Using the database of prostate histological images we have, we examined the first order statistical measures of the distributions of curvelet coefficients for each of the four pattern classes under consideration by pooling together coefficients in all directions of the image patches, their capabilities in class discrimination were studied and proven for feature selection in our classifier design. This is the first contribution.

The second major contribution of the thesis is the success of the design of a 2-level classifier with one support vector machine in each level for recognition of four classes of critical Gleason patterns in prostate cancer histological images. Two intermediate classes, Gleason 3+4 and Gleason 4+3, are correlated to the two classes, Class 1 (Gleason 3+3) and Class 2 (Gleason 4+4), we developed the first support vector machine with a voting mechanism by multiple

samples within a core image. Using the large number of these subimages of Class 1 and Class 2, the training success with unanimous voting assures the adequacy of the selected features used in the first support vector machine with maximum margin. Any non-unanimous voting result will assign the subimage into Class 3 (Gleason 3+4 and Gleason 4+3) which will be discriminated in the second support vector machine with a majority voting in level 2. This 3-class machine structure in level 1 is distinct from the standard multi-class support vector machines. The classifier was trained using 32 sample images of each class, each image was partitioned into 25 half-overlapped image patches with a moving window method to provide 1600 training samples for each class.

The designed classifier was successfully trained with 100% training accuracy. The 10% Jackknife validation result reaches an overall 97.97% accuracy; the first SVM gave an 98.88% accuracy for the classification of primary Gleason grade 3+3 and 4+4, which is outstanding as compared to the published studies [34, 35]. The validation result of the second SVM attained an accuracy of 95.58% which was about 3.58% higher than the only available reported work regarding the in-between classes [99]. The classifier was tested with 24 samples of each class and achieved an overall testing accuracy of 95.83%. This is the major contribution of this dissertation research.

5.1   SUGGESTIONS FOR FURTHER RESEARCH

The following two problems are suggested for further research:

1. The significant curvelet coefficients (the local max curvelet coefficients) that have been used here carry the unique orientation information at each location, and thus the orientation co-occurrence matrix can be evaluated in a straight forward manner. This will enable us to extract the second order statistical texture features reflecting some structural characteristics as well. This will be of great interest in further study on classification of critical prostate cancer tissue images.

2. The bound on the number of support vectors in a classifier is an important problem in training a support vector machine giving maximum margin and minimum error under a reasonable upper bound on support vector number, the smaller bound will assure the reliability of the trained machine. The analytical study toward a feasible of this optimization process is suggested for future study.

APPENDIX

A. CURVELET IMAGE DENOISING

The curvelet transform provides a sparse representation of image with rich curvilinear in multiscale and orientations which implies that most of the curvelet coefficients will have close to zero values except for those dominant curvelet coefficients localizing the edginess. The distribution of very few but large magnitude coefficients and the majority that has low amplitude coefficients contributes the heavy tailed shape distribution with peaks always occur at zero (leptokurtic). Figure 15 demonstrates such histograms representing curvelet coefficients in subband 4 and 5 obtained from P3S3 and P4S4 prostate sample images of 256×256 pixels size. The curvelet coefficients across all orientations are pulled together in each scale so that the distribution is nearly rotation invariant.

The multiscale Meyer wavelets provide an infrastructure that can be used to construct the narrow curvelet wedges systematically at different scales and to divide the 2-D Fourier transform in the sense of partition of unity. The needle-shape elements of this transform resulting from the multiscale directional wavelet are sharply supported in the spatial domain and are highly directional sensitive and anisotropic. In the transform domain, a curvelet coefficient with significant magnitude denotes a short edge segment centered at the location $k = (k_1, k_2)$, in orientation θ_l and at scale level j, as the curved singularity shown in Figure 14. The curvelet coefficients with large amplitude provide optimal sparse representations of edge components of images and information embedded therein can deliver useful description of the image for different types of image processing purposes, e.g. edge detection.

For image denoising task, the moderate white Gaussian noise of a corrupted image are often associated with weak magnitude coefficients at fine scales and one can recover the image by properly excluding most of the weak coefficients and reconstructing a preferable result with the rest fine and coarse scale coefficients.

A.1   CURVELET SHRINKAGE

Suppose the observed 2-D signal has been corrupted by additive noise and is given by

$$g = f + \varepsilon \tag{A.1}$$

where ε is white zero-mean Gaussian noise with variance σ^2 which is independent of the signal f. We observe the contaminated signal g and wish to recover the signal f containing as much as much important information as possible. Equivalently, the goal is to exclude the noise ε to obtain an estimate which minimizes the mean squared error (MSE)

$$MSE(\hat{f}) = E\left[\left(\hat{f} - f\right)^2\right]. \tag{A.2}$$

In the curvelet domain, the problem can be formulated as

$$y_{j,l} = c_{j,l} + n_{j,l} \tag{A.3}$$

where $y_{j,l}$ represents the noisy curvelet coefficients at scale j, orientation l, and the $c_{j,l}$ denotes the true coefficients, and $n_{j,l}$ is the independent Gaussian noise. The goal is to estimate $\hat{c}_{j,l}$ from the corrupted observation $y_{j,l}$. Because the computation of the variance σ^2 of the noise is prohibitively expensive, in the work of image denoising using curvelet transform by Starck et al. [16], they calculated an approximate value $\tilde{\sigma}_\lambda^2$ of individual variances with Monte-Carlo simulations and simply estimated the variance values by evaluating the curvelet transform of few standard white noise images. The λ is a set of parameters of the ridgelet (curvelet) transform. The true coefficients are then estimated with the hard shrinkage function

$$
\hat{c}_\lambda^{hard} = \begin{cases} 0, & \text{if} \quad |c_\lambda|/\sigma < k\tilde{\sigma}_\lambda \\[2em] c_\lambda, & \text{if} \quad |c_\lambda|/\sigma \geq k\tilde{\sigma}_\lambda. \end{cases} \tag{A.4}
$$

where k is chosen based on the scale in their experiments. The principle of the hard curvelet shrinkage is to shrink small curvelet coefficients toward zero to remove noise while retain the remaining few large ones that explain the useful information of images.

B. GLEASON GRADING FOR PROSTATE CARCINOMA

According to the epidemiology statistics [48], over 161,300 anticipated new cases of prostate carcinoma will be diagnosed and near 26,730 men are likely to die in the United States in 2017. Since 1966, the Gleason grading system devised by Donald Gleason has been established as a standard for interpreting this disease by expert pathologists based on microscopic tissue images from needle biopsies [49-50]. Gleason grade is categorized into 1 to 5, and increases based on the cumulative loss of regular glandular structure which reflects the increasing degree of malignancy aggressive phenotype, as illustrated in Figure 33. When an urologist examines a tissue image, he initially examines for a major (primary) arrangement of glandular architecture (an area where the adenocarcinoma acts to be most aggressive) and then inspects for a less extensive Gleason grade pattern (secondary) of growth, and assigns each one a Gleason grade number, primary grade and secondary grade respectively. The Gleason score (GS) is the summation of the primary grade and the secondary grade, ranging from 5 to 10. Urologic pathologists incline to determine a total score of 6 as a slower-growing cancer, 7 (3+4) as a medium-grade, and 4+3, 8, 9, or 10 as poorly differentiated (higher-grade) more aggressive carcinoma. Examples of TMA prostate images of Gleason grades P3S3, P3S4, P4S3 and P4S4 in Figure 34 are shown for the different complicated texture patterns from benign, critical intermediate class to carcinoma class.
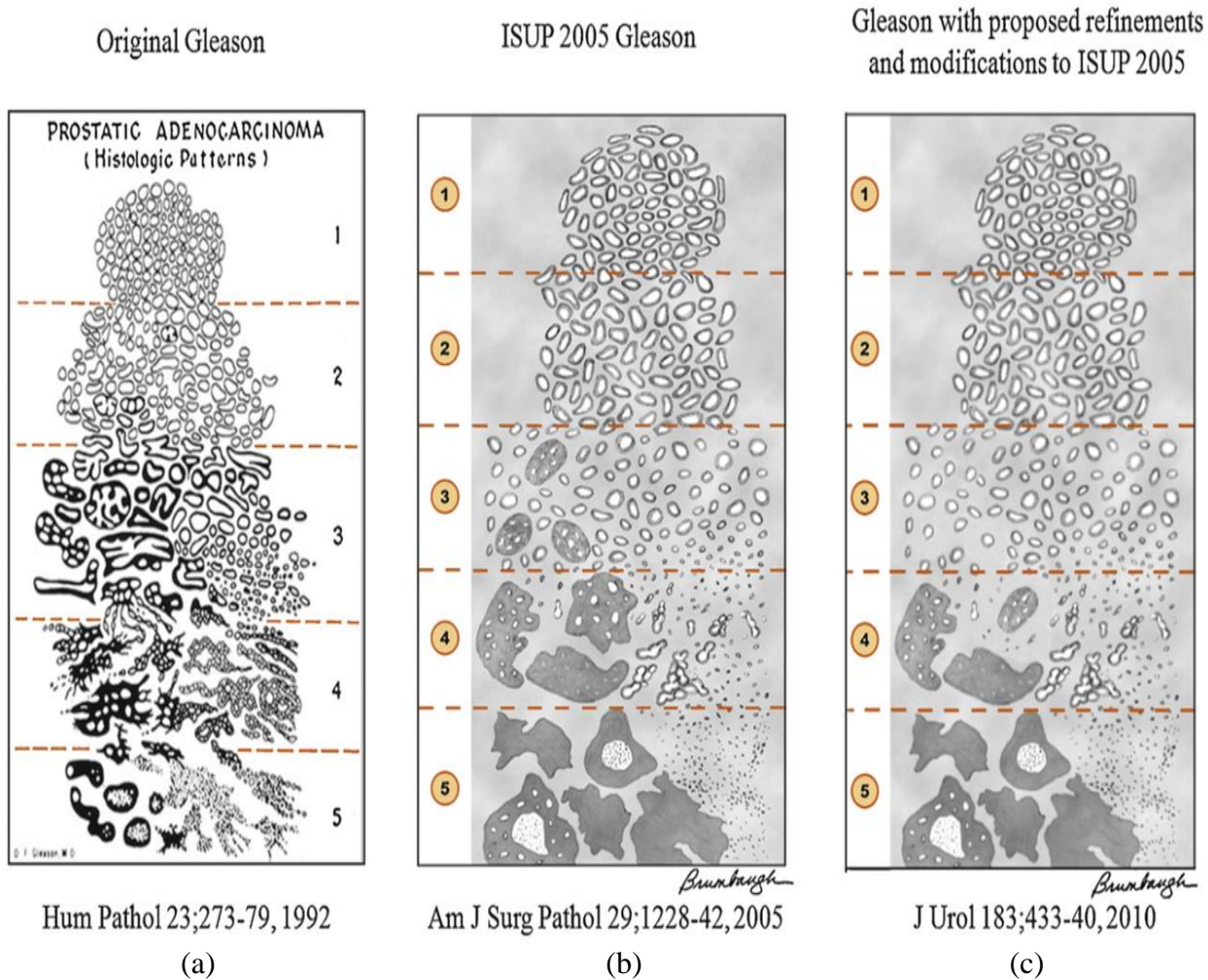
Figure 33. The Gleason grading system diagram (a) original by D.F. Gleason, M.D and (b) the modification made in 2005 and (c) the right hand side are the refinement made in 2010.
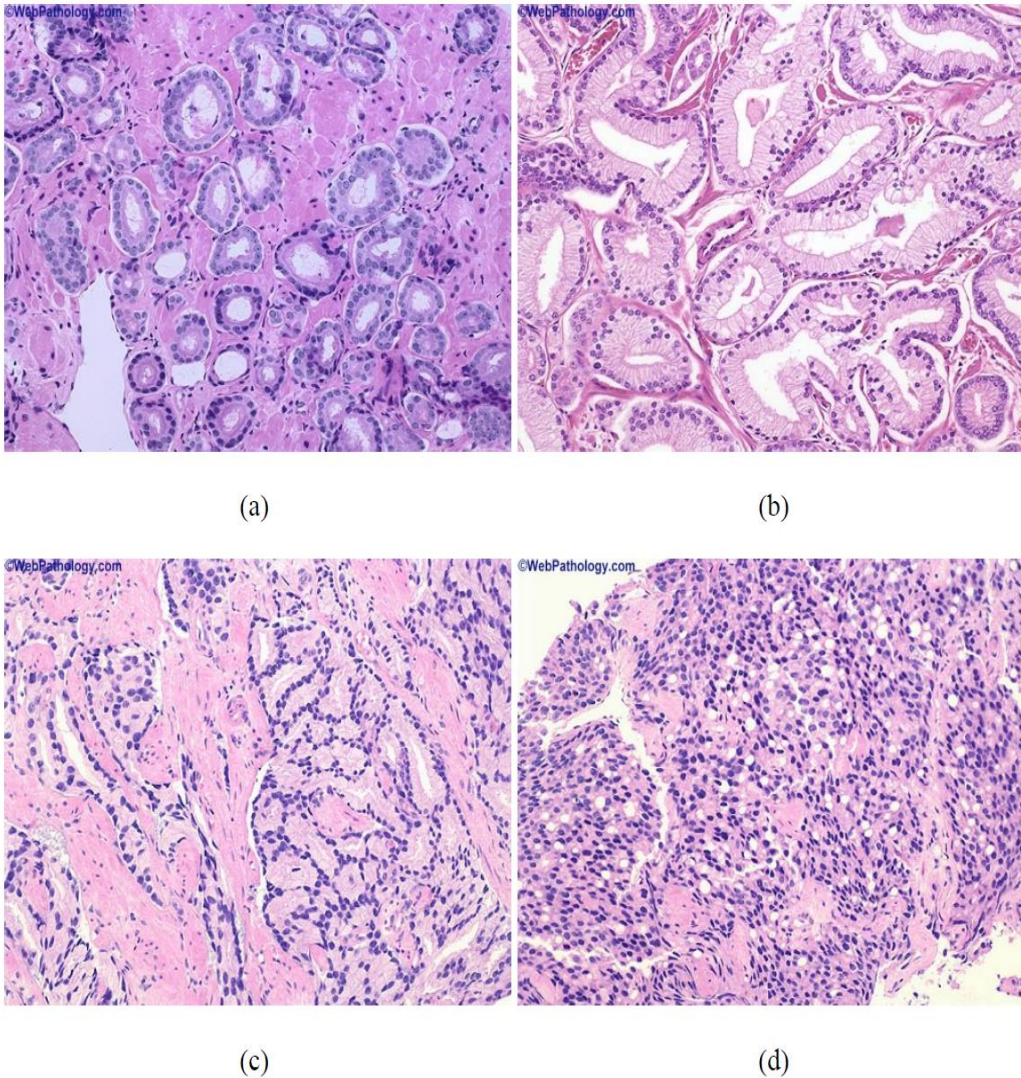
Figure 34. Examples of prostatic carcinoma (a) Gleason grade 2, (b) Gleason grade 3, (c) Gleason grade 4, and (d) Gleason grade 5, courtesy of WebPathology.com

A lower carcinoma grows more slowly and is less likely to propagate than a higher grade carcinoma [51]. The primary Gleason grades 3 and 4, hence, the Gleason Scores 6 and 7, are discriminated as the mid-point between the low-grade (less aggressive) and the intermediate-grade carcinoma and it generates the most lack of consensus in second-opinion assessments. In order to determine a patient's prognosis, there will be attention paid to more accurate Gleason grading and

scoring on all patients especially with a total Gleason Score of 7 (Primary 3 and Secondary 4 (P3S4), or Primary 4 and Secondary 3 (P4S3)), or 8 (Primary 4 and Secondary 4 (P4S4). The existence of a grade 4 is considered as a more aggressive regardless the overall degree of involvement [39, 49-50]. These scores (GS6, GS7, and GS8) are the critical Gleason scores that will help prognosis and provide suggestions for adequate definitive treatment. Examples of prostatic adenocarcinoma for the original Gleason grade 2 to 5 are shown in Figure 34.

Figure 34 (a) shows the stroma which is the fibromuscular tissue, clearly surrounds gland units, each of which is made of epithelial cells around a lumen; this retains to Gleason grade 2, where the glands are still round to olive-shaped and are nearly uniformly distributed. Figure 34 (b) is Gleason grade 3, where glands are shown with an irregular arrangement due to the progress of cancer. Figure 34(c) illustrates Gleason grade 4 whereas cancer gets worse, epithelium cells replicate irregularly and tend to occupy lumen areas. Figure 34 (d) shows Gleason grade 5, where lumens and stroma are virtually vanished.

Tabesh et al. [73] aggregated color, texture, and morphometric features at the global and histological object levels for classifications. They found that features associated with texture coarseness perform particularly well in these tasks. Doyle et al. [55] present a boosted Bayesian multiresolution (BBMR) system to identify regions of prostate cancer on digital biopsy slides. The algorithm decomposes the whole-slide image into an image pyramid comprising multiple resolution levels. Regions identified as cancer via a Bayesian classifier at lower resolution levels are subsequently examined in greater detail at higher resolution levels by a random forest classifier ensemble obtained by bagging multiple decision tree classifier. They concluded the classifier is able to detect areas involved by disease across multiple image resolutions which is similar to the approach employed manually by pathologists.

Kein and Jain [17] proposed a prostate cancer detection approach for the whole slide tissue image by extracting a cytological feature which presents the cancer nuclei (nuclei with prominent nucleoli) in the tissue, and apply this feature to detect the cancer regions. They showed the robustness of the cytological feature with the use of cytological-textural feature and conventional image texture features combination method. They later present another approach focused on the tissue structure and Gleason grading system specification. Their hierarchical classification scheme obtains 85.6% accuracy in classifying an input tissue pattern into one of three classes: benign, grade 3 carcinoma and grade 4 carcinoma.

Veltri et al. [59] proposed an adaptive active contour scheme (AdACM) that combines boundary and region based energy terms with a shape prior in a multi-level set formulation to segment nuclei on prostate cancer tissue microarray images. They achieved a classification accuracy of 84% in discriminating Gleason patterns 6, 7, 8, and 9 using a Support Vector Machine classifier and morphological features extracted from the segmentations.

Lopez and Again [35] introduced a new combination of wavelet and fractal features to automatically grade Gleason patterns 3, 4, and 5 pathological images using the Gleason grading system. They used pairwise coupling SVM classifiers for image classification and obtained a system accuracy close to 97%, estimated through three different cross-validation schemes.

The curvelet-based methodology for multiresolution texture feature extracted from prostate pathological images is described in next section. The exploration of using different strategies in ensemble classifier design is examined and reported. Non-linear support vector machines are employed as the baseline model of the tree structure classifier for classification of critical Gleason scores 6, 7 and 8. The strategy of ensemble classifier design and utilization of multiresolution texture features greatly improve the classification accuracy of prostate histological patterns.

C. PARAMETERS OF SVM WITH RBF FOR PROSTATE CANCER

CLASSIFICATION STUDY

Table 17. Parameters of the first SVM of the tree classifier for prostate cancer Gleason score 3+3 versus 4+4 (32 P3S3 and 32 P4S4, 25 patches per each image, 8 features are used, σ = 1.5545, *C* = 1).

Index	Support Vector								Alpha
1.	-0.0749	0.8678	-0.1143	-0.0165	0.0496	0.9969	1.2417	-0.3857	0.8249
2.	-0.6079	0.4502	0.1254	-0.0395	-0.0013	-0.2860	-0.2898	1.2007	0.0802
3.	-0.8761	0.1483	0.2083	0.4882	0.7888	-0.5470	-0.5376	-0.7697	0.7501
4.	-0.9215	1.8442	0.1451	0.3249	0.6276	-0.7946	-0.7006	-0.9808	0.0425
5.	-0.9597	0.1906	0.1602	-0.0019	0.0805	-0.7876	-0.6563	-0.2971	0.6054
6.	0.1566	0.3214	-0.1589	-0.0644	-0.0409	0.4038	0.6332	0.5377	0.6970
7.	-1.0388	-0.0125	0.2222	-0.0095	0.0153	-0.8039	-0.7287	0.0395	1.0000
8.	-1.1850	1.9580	0.2432	0.2039	0.0236	-0.9751	-0.8264	-0.9851	1.0000
9.	-0.4162	1.3639	-0.0941	0.2977	0.4845	0.2268	0.2995	-0.8458	0.4317
10.	0.0924	0.9940	-0.1262	-0.0138	0.0357	1.9508	3.2913	-0.2602	0.5735
11.	0.2670	0.8898	-0.1325	-0.0368	0.0159	1.8752	2.8340	0.5559	0.6345
12.	0.5121	0.9570	-0.1301	0.0012	0.0616	2.8451	4.3491	0.1789	1.0000
13.	0.0233	0.3849	-0.1159	-0.0153	0.0610	1.1901	1.4811	2.0255	0.8215
14.	-0.5153	0.0377	-0.0888	0.0052	0.0796	0.0463	0.0553	0.9636	0.3395
15.	-0.5254	0.2409	-0.1199	-0.0341	0.0300	-0.4591	-0.4172	0.6530	0.1227
16.	-0.9481	-0.1389	1.2287	0.9126	2.1221	-0.3291	-0.3792	-0.1084	0.8418
17.	-1.5584	-0.0044	38.2466	34.9549	36.0842	-1.0630	-0.8856	-1.0834	1.0000
18.	-0.7951	-0.1546	1.4612	0.9652	2.1713	0.1256	0.1119	1.2246	1.0000
19.	-1.2589	-0.9674	3.3732	6.2525	3.8682	-0.5798	-0.5563	-0.5255	1.0000
20.	-1.4381	-0.5505	8.7774	17.6508	15.4115	-0.9006	-0.7954	-1.0172	1.0000
21.	-0.2123	0.8491	-0.1486	-0.0811	-0.1159	-0.4849	-0.2319	0.7635	0.1302
22.	-0.1718	0.2548	0.0752	0.4558	0.2368	0.8790	0.8310	2.3112	0.7389
23.	-0.4508	-0.0298	0.0485	0.4468	0.2325	0.4661	0.4559	1.5601	0.3813
24.	1.2052	-1.1744	-0.1700	-0.1416	-0.1791	0.6696	0.5223	7.5350	0.2641
25.	1.6459	-0.9934	-0.2349	-0.1389	-0.1752	2.8690	3.7552	2.7198	0.3254
26.	5.9169	-0.9704	-0.2603	-0.1428	-0.1643	7.6255	9.8389	2.1384	0.5266
27.	4.1370	-0.9553	-0.2563	-0.1327	-0.1415	4.2869	4.9708	0.7630	0.3994
28.	0.2073	-0.9245	-0.1517	-0.1443	-0.1885	-0.4244	-0.3507	0.8984	0.4356

Table 17 (continued)

29.	1.1802	-0.9669	-0.1689	-0.1460	-0.1898	1.1986	1.7484	2.0510	0.2923
30.	1.2498	-1.2124	-0.1686	-0.1292	-0.1466	0.7061	0.3631	5.3433	0.3669
31.	1.5362	-1.2616	-0.2373	-0.1352	-0.1569	1.1270	0.6670	8.5108	0.3642
32.	2.3734	-1.1623	-0.2449	-0.1607	-0.1917	2.6646	2.6725	5.6584	0.1762
33.	2.0659	-1.2153	-0.2371	-0.1422	-0.1797	2.8334	2.6895	5.1440	0.1128
34.	2.5373	-1.0855	-0.2410	-0.1585	-0.1820	3.1430	2.8501	7.8427	0.4690
35.	2.8646	-1.0726	-0.2417	-0.1449	-0.1904	4.4334	4.6330	5.1519	0.3407
36.	-0.4143	-1.0705	-0.0954	-0.1241	-0.1645	-0.5631	-0.6418	0.5403	0.3807
37.	0.3373	-1.0640	-0.1132	-0.1237	-0.1580	0.8895	0.7542	1.3435	0.6472
38.	2.2444	-1.1259	-0.2326	-0.1342	-0.1833	4.0681	4.1548	2.6845	0.0831
39.	1.7535	-1.0533	-0.2283	-0.1288	-0.1705	3.6348	3.8500	1.5327	0.0388
40.	-0.5421	-0.7055	-0.0992	-0.0580	-0.0474	-0.7958	-0.7364	-0.6204	0.8378
41.	-0.1305	-1.2223	-0.1193	-0.1335	-0.1737	-0.4275	-0.5293	3.6564	0.3978
42.	2.1841	-1.2137	-0.2331	-0.1300	-0.1892	3.0321	2.6103	4.0427	0.1254
43.	3.5129	-1.2044	-0.2474	-0.1322	-0.1789	4.0310	3.6580	5.9644	0.2898
44.	0.2520	-0.8131	-0.1308	-0.0415	-0.1496	0.6705	0.7439	-0.4602	0.7403
45.	3.6991	-1.0801	-0.2458	-0.1380	-0.1814	6.0922	7.2741	3.9911	0.4929
46.	2.9698	-0.9715	-0.2404	-0.1328	-0.1710	6.3372	8.5473	1.5371	0.4173
47.	2.1546	-1.0396	-0.2261	-0.1315	-0.1709	4.9354	6.1492	1.8426	0.3724
48.	2.8323	-0.9735	-0.2386	-0.1348	-0.1745	6.0011	7.9817	1.8683	0.1026
49.	1.2857	-0.9590	-0.1862	-0.1391	-0.1726	-0.0934	0.0689	1.9637	0.0829
50.	1.9641	-0.8617	-0.2344	-0.1344	-0.1681	3.6516	3.9301	0.0454	0.0678
51.	2.1589	-0.9707	-0.2323	-0.1318	-0.1660	4.3515	4.9524	0.6761	0.2129
52.	1.6516	-0.9262	-0.2284	-0.1314	-0.1594	3.2569	3.4655	-0.1085	0.4164
53.	1.6470	-0.8417	-0.2299	-0.1282	-0.1521	2.1920	1.7536	-0.2641	0.1801
54.	0.3207	-1.1476	-0.1274	-0.1248	-0.1551	0.6288	0.5124	3.0152	0.3576
55.	3.5111	-1.0507	-0.2476	-0.1370	-0.1771	4.5988	4.7665	2.9254	0.1722
56.	-0.8179	-0.7902	0.1106	0.0022	-0.0207	-0.8559	-0.8062	-0.8551	0.9142
57.	0.8468	-1.1160	-0.1334	-0.1177	-0.1511	1.9814	2.2365	0.5282	0.3108
58.	3.2227	-1.0841	-0.1813	0.0638	0.3496	3.4269	2.4721	1.5612	0.2861
59.	4.1921	-1.1597	-0.2537	-0.1118	-0.1986	3.4734	4.2871	3.0839	0.3306
60.	1.7467	-1.2917	-0.1553	0.1041	0.3647	1.8231	0.8768	1.6411	0.0269
61.	-0.0224	-1.4829	0.6293	1.6684	0.9844	1.0124	0.4300	1.2045	0.6494
62.	0.8938	-1.4160	0.3135	0.2485	0.7513	2.1266	1.2818	3.8983	0.3958
63.	3.5766	-1.0421	-0.2452	-0.0906	-0.0617	2.3903	2.0637	2.5905	0.1436

Table 17 (continued)

64.	2.2383	-1.1436	-0.2428	-0.0874	-0.0627	0.3541	0.5621	1.0745	0.1158
65.	3.3859	-1.0533	-0.2455	-0.0995	-0.1966	1.6491	1.9008	4.5860	0.3223
66.	2.4122	-1.0334	-0.2525	-0.1425	-0.1794	0.9166	1.1077	3.7635	0.1885
67.	2.5746	-0.9141	-0.2502	-0.1381	-0.1744	1.4570	1.3209	0.0945	0.0453
68.	3.9034	-0.7929	-0.2603	-0.1434	-0.1699	1.6454	2.4440	0.1682	0.3742
69.	2.7996	-1.0403	-0.2553	-0.1372	-0.1611	0.9884	1.4151	0.5116	0.0535
70.	1.7153	-0.8559	-0.2488	-0.1366	-0.1666	0.1500	0.3046	-0.2051	0.1419
71.	2.0439	-0.9467	-0.2498	-0.1281	-0.1451	0.4378	0.6671	1.9123	0.0136
72.	3.5627	-1.0361	-0.2580	-0.1424	-0.2304	1.3053	2.1311	1.9003	0.2140

Bias
-0.4702

Table 18. Parameters of the Second SVM of the tree classifier for prostate cancer Gleason score 3+4 versus 4+3 (32 P3S4 and 32 P4S3, 25 patches per each image, 10 features are used, σ = 2.7974, *C* = 6.1212).

Index	Support Vector										Alpha
1.	-1.0346	1.3610	0.7598	1.5600	0.0584	-0.6666	-0.7772	-0.4761	-0.5528	-0.6369	6.1212
2.	-0.2347	0.4206	0.2121	-0.2943	-0.2213	-0.3267	-0.2675	-0.7985	-0.9298	1.0444	6.1212
3.	-1.2319	0.6686	1.1185	0.0191	2.3406	-0.6991	-0.6403	-0.8980	-1.1418	1.5724	1.8048
4.	-0.6498	0.3822	0.6882	-0.1079	-0.0063	-0.5115	-0.6336	-0.8031	-0.9541	0.6583	6.1212
5.	-0.8533	-0.2619	0.7349	-0.1559	-0.0339	-0.5529	-0.1607	-0.6454	-0.6131	1.0320	6.1212
6.	0.3403	-1.2454	-0.5910	-0.1471	0.2365	1.7283	1.0603	-0.6934	-0.7386	0.1729	1.6640
7.	-1.0516	-0.3035	1.3266	0.0692	0.0587	-0.5747	-0.4276	-0.6266	-0.5355	0.8606	6.1212
8.	-1.1705	0.6752	1.1356	-0.0317	0.0418	-0.6801	-0.6238	-0.7133	-0.6940	1.1121	6.1212
9.	-1.1671	0.6751	1.2140	1.6147	0.4752	-0.6445	-0.7494	-0.8942	-1.1154	1.8956	6.1212
10.	-0.5507	0.2642	0.4822	-0.3493	-0.2122	-0.3750	-0.0472	2.1883	2.1399	-1.5862	3.2852
11.	-1.0877	0.6726	0.8929	-0.2958	-0.3905	-0.6428	-0.3314	-0.5402	-0.4745	0.8536	4.3390
12.	-0.6149	-0.2555	0.6403	-0.1682	-0.1079	-0.4864	-0.1559	0.4354	0.5481	-0.9735	2.0786
13.	-0.7623	-0.9685	0.7881	-0.1672	-0.0174	-0.4807	0.1500	-0.5037	-0.4091	0.3811	6.1212
14.	-0.4675	-1.1497	1.3698	0.2522	0.1722	0.1362	-0.0099	1.4995	1.5426	-1.3236	4.8477
15.	0.3727	-0.3608	-0.5043	2.1477	4.4105	-0.0293	0.0882	0.8915	1.0439	-1.6821	0.1721
16.	0.0231	0.7376	-0.7957	0.8403	2.0411	-0.4487	-0.4576	0.0037	0.2117	-0.9853	2.5017
17.	-0.9212	0.5045	0.8557	2.2864	4.0651	-0.6286	-0.7932	-0.7906	-0.9777	-0.4402	3.7260
18.	1.0353	-0.0881	-1.0916	-0.0293	0.5954	0.2958	-0.1948	-0.4942	-0.4749	-0.4024	6.1065
19.	0.1483	0.7692	-1.0403	-0.4111	-1.2963	-0.1680	0.4610	3.3592	2.8931	-2.1424	1.1156
20.	3.1264	1.6241	-1.1780	-0.2580	-0.2544	4.0516	0.0178	-0.3838	-0.1982	0.3820	1.0071
21.	0.1033	-0.0701	-0.9666	0.0386	0.1181	-0.1961	-0.2673	-0.6096	-0.6220	-0.1106	6.1212
22.	1.4099	0.4006	-1.2071	-0.3682	-0.2661	0.6488	0.7399	2.0685	1.8430	-1.1011	0.5261
23.	2.7368	-0.0921	-1.2855	-0.6192	0.0702	2.5277	1.2456	0.2050	0.4584	-0.0409	3.1314
24.	0.8839	-0.6003	-1.2042	-0.4188	-0.3731	0.0160	0.9088	0.5710	0.7063	-0.8487	1.0103
25.	-0.0382	0.4422	-0.9178	-0.1670	0.4830	-0.2889	-0.4663	-0.7402	-0.8499	0.2515	1.2961
26.	0.9492	-0.6798	-1.2251	-0.4876	-0.4082	-0.0695	0.9642	0.9483	0.9769	-1.0718	6.1212
27.	0.9706	-0.6029	-1.2244	-0.4893	-0.4184	0.0381	0.8472	0.4020	0.5381	-0.7010	6.1212
28.	3.1562	-0.5608	-1.1945	-0.7571	-1.4366	4.3944	3.8555	1.2650	1.3390	-0.8457	1.3263
29.	1.8841	-0.5711	-1.0449	-0.7534	-0.5970	2.1574	3.7166	-0.5789	-0.5829	0.7253	2.9944
30.	1.8142	0.0958	-1.0366	-0.3608	-0.4605	1.7722	0.5037	0.5942	0.6940	-0.6280	2.8570
31.	1.2560	-0.6743	-0.9841	-0.3490	-1.3983	1.2628	2.6516	1.0845	1.3773	-0.8440	0.5960
32.	-0.5742	0.4585	0.3866	-0.3963	-0.4995	-0.5143	-0.3478	-0.2765	-0.3893	-0.2977	6.1212
33.	-1.2434	0.5009	1.9527	3.0944	4.1699	-0.6230	-0.8088	-0.9229	-1.2051	1.9049	1.4181

Table 18 (continued)

34.	0.1073	-1.0380	-0.4657	2.0537	4.3489	-0.0650	0.7989	2.1335	2.3125	-2.0090	1.1960
35.	-1.1790	-1.7970	7.0000	6.7277	14.5749	-0.5889	0.0787	-0.1647	-0.0718	-2.1767	1.1313
36.	-0.7100	-1.0445	1.7302	1.7794	3.8713	-0.5044	0.9799	-0.0455	0.1410	-0.9066	0.3771
37.	0.2351	-0.0244	-1.0590	-0.3118	-0.3690	-0.1892	-0.1062	0.7175	0.8657	-0.9326	6.1212
38.	-1.7164	1.4574	5.7226	2.9177	4.1002	-0.7704	-0.9211	-0.8682	-1.1230	-0.7525	0.7810
39.	-1.7344	2.3146	5.4891	2.7276	3.7507	-0.7856	-0.9132	-0.8903	-1.1498	-0.0327	0.2724
40.	0.0098	-0.1918	-0.9724	-0.3368	-0.3406	-0.2062	0.1679	-0.3198	-0.3153	-0.3049	1.4378
41.	-0.7034	0.8496	0.4560	-0.2762	-0.4461	-0.5968	-0.5380	-0.0193	-0.0251	-0.5972	6.1212
42.	-0.5216	0.7309	0.3842	-0.4272	-0.5615	-0.3874	-0.4023	-0.7600	-0.9197	1.4394	6.1212
43.	4.4544	0.2689	-1.3181	-0.8223	-1.4780	4.1183	2.9881	-0.0405	0.2162	0.2780	1.0677
44.	2.0397	-0.4044	-1.2276	-0.8001	-0.5152	1.3188	1.8622	0.6950	0.9224	-0.4609	0.9147
45.	1.9858	-0.4604	-0.9283	0.1551	0.0676	1.0564	2.9196	6.4765	5.4171	-2.1117	0.8612
46.	1.8499	0.2023	-1.1863	-0.4223	-0.7769	0.8887	1.5041	3.6187	3.4360	-1.4490	2.1921
47.	1.0071	-1.0197	-0.2629	2.6606	1.0477	1.1413	2.5863	8.0695	6.3827	-2.7809	0.8512
48.	1.5898	-0.0333	-1.0524	0.4740	-0.3546	0.9928	2.0501	9.0337	7.5775	-2.2651	0.7881
49.	2.4476	0.6509	-1.3378	-0.7773	-1.3125	1.0434	1.4335	6.2089	5.3192	-1.8348	1.0617
50.	1.1103	0.6956	-1.1605	-0.4029	-0.5975	0.2379	0.1102	0.1150	0.2848	0.1828	6.1212
51.	-0.4155	-0.4790	0.2719	-0.3388	-0.5194	-0.4028	0.2781	-0.3184	-0.3304	-0.2051	3.8364
52.	-0.2484	-0.6228	0.2077	-0.4057	-0.6162	-0.3305	0.4019	-0.1566	-0.1532	-0.0404	6.1212
53.	-0.6809	-0.3179	0.5489	-0.0669	0.2460	-0.5092	-0.3281	-0.5840	-0.5437	0.1952	6.1212
54.	-0.0998	-1.7050	-0.6373	0.0331	0.3064	0.9204	3.4898	-0.5520	-0.4806	-0.0037	1.5908
55.	-0.3101	-2.2455	1.0241	0.1784	0.3296	0.9905	3.5262	-0.3653	-0.2759	-0.2858	1.0102
56.	3.3198	0.5064	-1.0189	-0.1648	0.0471	6.4929	1.1396	1.3669	1.3091	-1.2056	0.7047
57.	-0.8093	0.3865	0.8806	1.6999	0.4135	-0.5162	-0.6880	-0.8352	-1.0263	0.4761	6.1212
58.	0.0217	-0.1021	-0.7633	-0.1308	-0.0472	0.3430	-0.2002	-0.0532	0.0849	-0.4248	6.1212
59.	-0.4233	-1.1208	0.7091	-0.1018	0.1527	-0.1997	0.1499	-0.7839	-0.9046	0.6319	6.1212
60.	0.3021	-0.4072	-0.7933	-0.0970	0.0270	0.4485	-0.1875	0.0065	0.0311	-0.9352	6.1212
61.	-1.0048	0.5055	1.0726	1.7324	0.3519	-0.5634	-0.5676	-0.5158	-0.4861	0.0468	6.1212
62.	-0.3537	-0.5255	0.3837	-0.2216	-0.1431	-0.2009	-0.0831	0.1533	0.2040	-0.8539	6.1212
63.	-0.4911	-0.8330	0.5222	-0.0545	0.2479	-0.3130	-0.2335	-0.0323	0.0068	-0.7343	6.1212
64.	1.4319	-0.3735	-0.8756	-0.3096	-0.3593	5.1708	1.4301	-0.1384	0.0220	0.2504	4.6796
65.	-0.9183	0.0131	1.0026	1.8841	0.7558	-0.5786	-0.6444	0.3101	0.3548	-1.6695	3.0772
66.	0.5961	-0.7298	-0.5468	1.3304	-0.0834	0.7884	1.0057	2.4946	2.3225	-2.1448	1.5940
67.	0.2056	-0.4735	-1.0285	-0.2202	-0.2153	-0.2319	0.4444	1.1582	1.2251	-1.3603	0.4046
68.	0.2153	0.2704	-0.9509	-0.2189	-0.0758	-0.1561	0.0092	0.3059	0.4168	-0.2251	6.1212

Table 18 (continued)

69.	-0.9990	0.8243	1.0063	1.9590	1.0372	-0.6180	-0.8016	-0.7508	-0.8447	0.0862	1.0048
70.	-0.4581	0.5859	0.4356	-0.0382	0.3417	-0.4474	-0.5646	-0.7567	-0.8330	0.8841	6.1212
71.	-0.5826	2.4840	0.6262	1.5646	2.5969	-0.3944	-0.8003	-0.8769	-1.0907	0.5994	0.7167
72.	-0.3252	0.2457	0.4187	-0.0381	0.4618	-0.3203	-0.4961	-0.4682	-0.4070	-0.1629	2.8991
73.	-0.7403	0.3845	0.9964	1.9397	0.8962	-0.4266	-0.5506	-0.7680	-0.8486	0.3084	6.0110
74.	-0.9952	0.1688	1.1095	2.1639	0.9219	-0.6205	-0.6574	-0.7651	-0.8530	0.2827	6.1212
75.	-1.0631	1.7481	0.9624	1.5135	0.2397	-0.6199	-0.7919	-0.8360	-1.0247	0.4867	0.2582
76.	-0.6522	-1.0439	0.8450	0.0339	0.3904	-0.4300	-0.2835	-0.6192	-0.7053	-0.8539	0.9313
77.	-1.0050	0.1599	0.8843	1.5985	0.7637	-0.6409	-0.5910	-0.8032	-0.9572	-0.0827	6.1212
78.	-0.9852	-0.2202	1.0128	0.0757	0.6172	-0.5690	-0.5640	-0.8004	-0.9105	0.3735	4.1702
79.	-0.5687	0.0210	0.1984	-0.4433	-0.5253	-0.5941	-0.0275	-0.1831	-0.0824	0.7146	1.7996
80.	-0.6717	0.1326	0.1306	-0.3877	-0.3141	-0.6347	-0.4469	-0.3206	-0.3001	-0.0133	4.7929
81.	-0.5557	-1.0698	0.5607	-0.2743	-0.2306	-0.4661	-0.1782	-0.4005	-0.4098	-0.2142	1.5337
82.	-1.1818	-0.3763	0.9260	-0.2699	-0.2673	-0.5723	-0.6106	0.1656	0.1463	-0.4300	0.2520
83.	0.6882	0.1965	-1.1477	-0.3771	-0.4464	0.2011	-0.3134	0.0707	0.2151	-0.3810	6.1212
84.	-0.9139	0.3543	0.7510	-0.3637	-0.4764	-0.3462	-0.6072	-0.5422	-0.6599	0.7043	6.1212
85.	-0.1892	0.0723	0.3381	1.2317	-0.0053	0.1591	-0.5688	-0.7975	-0.9383	0.9483	0.1632
86.	-1.4888	-0.0514	1.2842	1.0156	-0.1456	-0.7634	-0.6696	-0.7271	-0.8143	1.2593	1.1246
87.	-0.3834	-0.6590	0.7177	-0.1943	-0.1523	0.2618	-0.5004	-0.8054	-0.9679	0.8964	6.1212
88.	-1.4639	-0.2093	1.1015	-0.2666	-0.2377	-0.7672	-0.7043	-0.7700	-0.8965	1.1037	6.0097
89.	-1.2873	1.0888	0.8440	0.8535	-0.2627	-0.7256	-0.7474	-0.8440	-1.0162	1.7367	6.1212
90.	-1.0510	0.8098	0.6625	-0.3866	-0.4700	-0.6822	-0.6635	-0.4889	-0.5791	0.2345	6.1212
91.	0.2273	-0.6347	0.1386	-0.3688	-0.2378	0.2769	-0.3485	-0.5556	-0.6406	-0.1089	0.2604
92.	-1.2087	0.8432	0.9686	1.2583	-0.0333	-0.7283	-0.7262	-0.7808	-0.9401	0.3128	6.1212
93.	-1.2922	-0.5897	0.9889	1.1466	1.7567	-0.7201	-0.6261	-0.7964	-0.9749	0.5635	5.9037
94.	-0.5097	0.6522	0.6676	1.1980	0.2653	-0.4829	-0.6791	-0.3237	-0.2879	0.1951	6.1212
95.	0.6866	-1.8567	0.2334	3.7539	6.8966	0.4234	0.1276	2.1646	2.2248	-1.3088	1.0573
96.	-0.9789	-2.9634	10.4121	24.3916	15.5304	-0.2698	-0.2654	0.7330	0.7071	-2.3807	0.8703
97.	-0.4288	-2.5298	5.4784	4.2719	9.0279	-0.4418	0.3292	1.6218	1.9489	-0.9565	0.7262
98.	1.4293	-1.7826	-0.2944	2.3745	0.3944	0.6175	0.7646	2.9864	3.3341	-1.0332	1.0586
99.	0.0409	-2.3140	3.6378	3.4540	6.5427	-0.1256	0.5481	2.1489	2.4179	-0.9974	0.5614
100.	1.4357	-0.3133	-1.2734	-0.4184	-1.4138	0.2253	0.8136	2.2171	2.2447	-0.5511	2.8874
101.	1.2845	-1.5597	-0.9965	-0.1881	-0.6248	0.3268	1.4940	5.6791	5.5913	-1.0726	0.2631
102.	-0.1029	0.0734	0.2014	-0.3310	-0.2332	-0.2898	-0.4855	-0.3533	-0.3322	0.1055	4.4600
103.	-0.4136	-0.0044	0.2923	-0.2530	-0.1689	-0.4170	-0.4800	-0.4183	-0.3670	0.1106	6.1212

Table 18 (continued)

104.	1.6977	0.2899	-1.3507	-0.4268	0.4302	0.2339	-0.3392	0.0864	0.3276	-0.3909	4.2583
105.	0.9064	0.3633	-1.2090	-0.3962	0.2613	0.2285	-0.3768	0.6796	0.9348	-0.6315	6.1212
106.	0.8839	-1.0570	-1.1579	-0.5065	-0.2209	0.8731	0.7725	2.6946	2.6546	-1.2089	0.6571
107.	2.3794	-0.5190	-1.4000	-0.6516	-0.6038	1.4416	0.6125	0.4809	0.6777	-0.2139	0.5952
108.	0.5775	-0.1543	-1.1976	-0.5261	-0.3392	-0.0155	-0.1915	0.8874	0.9360	-0.8341	3.2673
109.	-0.0926	0.0177	0.0193	-0.3848	-0.1817	-0.4187	-0.3055	-0.5631	-0.5580	0.6204	6.1212
110.	-0.2943	0.1965	0.0414	-0.3360	-0.0309	-0.4890	-0.3780	0.0445	0.0018	-0.6412	6.1212
111.	-0.6572	-0.4035	0.3878	-0.2360	1.4550	-0.5576	-0.5562	-0.8375	-1.0368	1.4844	1.8734
112.	-0.0392	-0.4152	0.1134	-0.4519	-0.5208	-0.2234	0.2646	-0.1899	-0.1445	0.4047	6.1212
113.	1.4623	-1.4147	-1.0695	-0.5250	-0.6283	2.3168	3.1436	0.1932	0.3658	0.2479	3.4635
114.	2.1852	-0.4846	-1.2431	-0.5993	-0.8017	1.9830	1.6255	0.0722	0.2901	0.9011	0.3049
115.	5.2361	-1.4222	-1.2865	-0.8685	-0.7063	7.9888	4.1018	0.2377	0.5179	0.6043	0.6867
116.	3.0774	-1.7916	-1.1611	-0.8356	-0.6556	5.2942	3.2751	-0.0679	0.1813	0.7940	0.7434
117.	0.4781	-0.4988	-1.1429	-0.5548	-0.7613	-0.0697	0.9813	0.0242	0.1299	0.4940	3.2252
118.	-0.0455	-0.6794	-0.0912	-0.4218	-0.2298	-0.4379	0.4783	2.0811	1.9896	-1.4135	6.1212
119.	0.1593	-0.9866	-0.0514	-0.3515	0.2716	0.0129	-0.0971	0.4237	0.4763	-1.1182	4.4581
120.	1.6934	-2.0440	-1.2284	-0.8227	-0.5992	1.0570	5.6784	0.1904	0.5239	0.7255	0.0210
121.	-0.0027	-0.3244	-0.0472	-0.5257	-0.6343	-0.3519	0.9361	0.1601	0.4037	1.0291	1.1313
122.	0.9868	-0.6011	-1.1626	-0.4348	-0.5280	0.4636	0.6601	0.5551	0.7638	-0.2745	3.4883
123.	-0.7135	-0.4892	0.8203	1.6189	0.2141	-0.4653	-0.4371	-0.6180	-0.6778	0.0947	6.1212
124.	-1.0145	0.2738	1.0067	1.7816	0.2484	-0.6306	-0.7461	-0.6585	-0.7057	0.6054	6.1212
125.	-0.6423	0.5799	0.8112	1.7646	0.4113	-0.2798	-0.7580	-0.6106	-0.6377	0.5005	6.1212
126.	-0.8842	-0.4920	0.7434	-0.0677	0.1572	-0.4761	-0.4521	-0.6237	-0.6345	0.8044	6.1212
127.	-1.0969	0.4774	0.7917	1.6322	1.7608	-0.6939	-0.8425	-0.8656	-1.0398	2.8741	1.4394
128.	-0.6548	-1.0404	0.5191	-0.1118	0.1257	-0.5378	-0.1806	0.1126	0.1403	-0.6086	3.1814
129.	0.1423	-1.4240	0.1796	-0.2002	-0.0596	-0.1388	0.7928	-0.2023	-0.1802	-0.4844	6.1212
130.	0.6051	-1.0781	-1.0351	-0.2909	-0.0518	0.2098	0.8534	-0.1527	-0.0800	-0.2365	6.1212
131.	1.2840	-0.4439	-1.1032	-0.3400	-0.1446	1.3650	0.3434	0.2602	0.3832	-0.2034	6.1212
132.	-1.1171	0.7429	0.7059	1.3665	0.3154	-0.6988	-0.8527	-0.6231	-0.7399	-0.0021	6.1212
133.	-0.8813	0.4409	0.5817	1.1333	0.2583	-0.5494	-0.7476	0.0021	-0.0680	-0.9435	6.1212
134.	-0.7566	-0.3721	0.5902	-0.2997	-0.1792	-0.4719	-0.5383	-0.0362	-0.1132	-0.5902	6.1212
135.	-0.8452	-0.7526	0.7360	1.2032	0.1711	-0.5656	-0.4061	1.5448	1.4964	-1.3525	4.6682
136.	-0.2154	-0.9030	0.5828	-0.3220	-0.3201	-0.1095	0.3305	4.0004	3.6780	-0.9683	0.0115
137.	-1.0768	0.7282	0.6749	1.3853	0.3017	-0.6587	-0.8417	-0.6660	-0.7630	0.4353	1.3598
138.	-0.1609	-1.6187	0.3508	-0.2113	-1.2380	-0.1568	0.9661	6.9018	6.0059	-1.8409	0.4890

Table 18 (continued)

139.	-0.8104	0.8099	0.4921	1.3010	1.8988	-0.5975	-0.8488	-0.8950	-1.1308	3.0762	0.3001
140.	-0.9191	-0.7737	0.6936	1.5714	0.4982	-0.6557	-0.6325	0.6705	0.6779	-1.4252	0.4854
141.	-0.6095	-1.6605	0.7598	-0.1398	-0.1304	-0.2834	0.6406	5.7927	4.9670	-1.4850	0.1617
142.	0.2369	-0.3109	-0.2042	-0.5080	-0.6049	-0.3263	0.0313	0.1895	0.0076	-0.3585	1.8151
143.	-0.0013	-0.5085	0.0960	-0.4438	-0.4741	-0.1069	0.3633	3.6013	3.3341	-1.0058	0.1889
144.	-1.6869	-0.9722	6.3476	3.7363	2.6016	-0.7648	-0.8803	-0.9238	-1.2160	2.0822	1.0806
145.	-1.6516	0.6633	5.9014	10.0951	3.4068	-0.7472	-0.9606	-0.9269	-1.2437	1.3438	0.8675
146.	0.2022	0.5617	0.6057	2.1650	2.3204	-0.2198	-0.7841	-0.7781	-0.9373	0.1314	4.1313
147.	-0.6853	-0.0407	0.8468	2.7403	2.8855	-0.5423	-0.8779	-0.8722	-1.1208	0.5179	2.8404
148.	0.9195	-0.3980	-1.0565	1.6839	2.4170	0.4592	-0.5740	-0.7755	-0.9531	-0.1676	0.4883
149.	1.2162	-1.0680	-1.2651	-0.3203	-1.2830	0.1341	1.2407	5.3923	5.0984	-1.4402	1.6206
150.	2.0575	-1.1231	-1.3699	-0.5531	-1.4689	1.6562	0.8370	2.9756	3.1240	-0.6832	0.4537
151.	0.4670	-0.6405	-1.2041	-0.2199	0.2687	-0.2030	-0.3889	-0.0705	-0.0512	-1.1326	6.1212
152.	2.2283	0.0623	-1.3286	-0.5674	-0.6114	2.1206	0.0018	0.0578	0.1133	0.6371	0.1929
153.	-0.2217	0.3628	-0.1251	-0.3150	-0.2816	-0.5463	-0.6253	-0.4440	-0.5451	-0.1327	6.1212
154.	-0.2710	0.5606	0.0565	-0.2360	-0.0925	-0.4595	-0.6763	-0.7912	-0.9533	1.8971	1.1672
155.	-0.6881	-2.5417	0.3583	-0.2121	-0.2061	-0.5873	1.3637	-0.5404	-0.5724	0.5544	1.3441
156.	0.5837	-1.9606	-1.1686	-0.4559	-0.4727	-0.0795	2.6835	1.0695	1.1914	-0.2740	0.1330
157.	-0.9529	-0.7510	0.5199	-0.1638	1.7628	-0.6450	-0.6161	-0.7316	-0.8929	0.0352	2.1541
158.	-0.7291	0.5288	0.2668	1.0463	1.5798	-0.6236	-0.7084	-0.7489	-0.9466	0.0444	6.1212
159.	1.0751	-3.3369	-1.0875	-0.3971	-0.6465	1.1053	12.9512	0.0094	0.1671	0.3188	0.8099
160.	1.4250	-3.1106	-1.0832	-0.7494	-0.7037	1.8202	17.3703	0.1041	0.3553	0.6018	0.8107
161.	0.3607	-2.8879	0.2096	-0.2750	-0.4626	0.3315	4.6444	-0.4928	-0.3888	1.0511	1.2187
162.	0.8037	-2.8037	-1.0179	-0.3248	-0.5013	1.1476	5.2245	-0.4604	-0.2950	1.3363	0.2257
163.	2.9145	-1.2026	-1.3501	-0.9133	-1.6647	1.8812	5.9897	1.3722	1.8943	0.5100	0.6664
164.	3.2989	-1.3750	-1.3937	-0.9067	-1.6658	2.3587	2.9275	0.8291	1.0956	0.1104	1.1618
165.	6.0310	-0.7655	-1.3662	-0.5308	-0.9221	8.3308	0.8997	-0.4066	-0.2339	1.6504	0.4400
166.	4.1520	-0.3426	-1.2359	-0.3372	-0.6863	8.1608	-0.2298	-0.6275	-0.5696	2.2271	0.4395
167.	5.0839	-1.6732	-1.4154	-0.8620	-0.9548	6.3622	3.9631	-0.5411	-0.3967	3.6246	0.5267
168.	3.2333	-1.7501	-1.3698	-0.8292	-0.8418	2.5127	4.3090	-0.3198	-0.0223	2.1360	0.5834
169.	-0.3298	-2.0812	0.5304	-0.2220	0.0581	-0.2323	0.8364	-0.0456	0.0803	-0.2976	0.7638
170.	2.3146	-0.4207	-1.1563	-0.4372	-0.3336	4.6665	1.0096	-0.0532	0.2119	0.9902	5.7482
171.	-0.6333	0.0862	0.3308	-0.3296	-0.0256	-0.4694	-0.4366	-0.8424	-1.0134	1.7604	6.1212

Bias
-0.1302

D. RANKING OF THE STATISTICAL FEATURES

Table 19. Ranking of the statistical features for the two classification problems

	GS 3+3 vs GS 4+4		GS 3+3 vs GS 4+4	
	Feature	**Subband**	**Feature**	**Subband**
1.	**Entropy**	**4**	**Entropy**	**4**
2.	**Skewness**	**4**	**Energy**	**4**
3.	**Energy**	**4**	**Skewness**	**4**
4.	**Skewness**	**5**	Variance	4
5.	**Energy**	**5**	Kurtosis	5
6.	Kurtosis	4	Entropy	5
7.	**Variance**	**3**	**Skewness**	**5**
8.	**Kurtosis**	**3**	**Energy**	**5**
9.	Variance	5	**Kurtosis**	**3**
10.	Entropy	5	**Variance**	**3**
11.	Kurtosis	5	Variance	5
12.	Skewness	3	Kurtosis	4
13.	Variance	4	Skewness	3
14.	Entropy	3	Energy	3
15.	Energy	3	Entropy	3

BIBLIOGRAPHY

[1]   M. Tuceryan and A. K. Jain, "Texture Analysis," In The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang (eds.), pp. 207-248, World Scientific Publishing Co., 1998.

[2]   Rao, C. V., J. Malleswara Rao, D. S. Jain, and V. K. Dadhwal. "Satellite image fusion using Fast Discrete Curvelet Transforms." In Advance Computing Conference (IACC), 2014 IEEE International, pp. 952-957. IEEE, 2014.

[3]   De Hoop, M. V., H. Smith, G. Uhlmann, and R. D. Van der Hilst. "Seismic imaging with the generalized Radon transform: a curvelet transform perspective." Inverse Problems 25, no. 2 (2009): 025005.

[4]   C. M. Lopez and S. Again, "A New Set of Wavelet- and Fractals-based Features for Gleason Grading of Prostate Cancer Histolopathology Images," Proc. SPIE-IS&T Electronic Imaging, v. 8655, 2013.

[5]   S. Mallat, "A Wavelet Tour of Signal Processing: the Sparse Way", 3rd ed., Chap. 5, 2009.

[6]   C. H. Chen and G. G. Lee, "On Multiresolution Wavelet Algorithm using Gaussian Markov Random Field Models, in Handbook of Pattern Recognition, Computer Vision, 2nd ed., Eds., C. H. Chen. L. F. Pau and P.S.P. Wang, Chap. 1.5., World Scientific, 1999.

[7]   K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation using Gabor Filters," Pattern Recog., vol. 34, pp. 1167-1186, 1991.

[8]   L. Dettori and A. I. Zayed, "Texture Identification of Tissues using Directional Wavelet, Ridgelet and Curvelet Transforms," in Frames and Operator Theory in Image and Signal Processing, ed., D. R. Larson, et al., Amer. Math. Soc., pp. 89-118, 2008.

[9]   L. Dettori and L. Semler, "A comparison of Wavelet Ridgelet and Curvelet Texture Classification Algorithms in Computed tomography," Computer Biology & Medicine, vol. 37, pp. 486-498, 2007.

[10]   G. Castellaro, L. Bonilha, L. M. Li and F. Cendes, "Multiresolution Analysis using wavelet, ridgelet and curvelet Transforms for Medical Image Segmentation," Intern. J. Biomed. Imaging, v. 2011, Article ID 136034, 2011.

[11] E. J. Candes and D. L. Donoho, "New Tight Frames of Curvelets and Optimal Representation of Objects with Piecewise Singularities," Commun. Pure Appl. Math., vol. 57, no. 2, pp. 219-266, 2004.

[12] E. J. Candes, L. Demanet, D. L. Donoho, and L. Ying, "Fast Discrete Curvelet Transform," Multiscale Modeling & Simulations, vol. 5, no. 3, pp. 861-899, 2006.

[13] E. J. Candes and D. L. Donoho, "Continuous Curvelet Transform: II. Discretization and Frames," Appl. Comput. Harmon. Anal., vol. 19, pp. 198-222, 2005.

[14] Donoho, David L., and Mark R. Duncan. "Digital curvelet transform: strategy, implementation, and experiments." In AeroSense 2000, pp. 12-30. International Society for Optics and Photonics, 2000.

[15] E. J. Candes, L. Demanet, D. L. Donoho, and L. Ying, "Curvelab Toolbox, version 2.0," CIT, 2005.

[16] J-L. Starck, E. Candes and D. L. Donoho, "The Curvelet Transform for Image Denoising," IEEE Trans. IP, vol. 11, pp. 131-141, 2002.

[17] K. Nguyen, A. K. Jain and B. Sabata, "Prostate Cancer Detection: Fusion of Cytological and Textural Features," Jour. Pathology Informatics, vol. 2, 2011.

[18] L. Guo, M. Dai and M. Zhu, Multifocus Color Image Fusion based on Quaternion Curvelet Transform," Optic Express, vol. 20, pp. 18846-18860, 2012.

[19] Wen-Chyi Lin, Ching-Chung Li, Christhunesa S. Christudass, Jonathan I. Epstein and Robert W. Veltri, "Curvelet-Based Classification of Prostate Cancer Histological Images of Critical Gleason Scores," In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pp. 1020-1023, 2015.

[20] Jianwei Ma and Gerlind Plonka. "The curvelet transform." Signal Processing Magazine, IEEE 27.2, pp.118-133 2010.

[21] J-C. Starck, F. Murtagh and J. M. Fadili, "Sparse Image and Signal Processing", Cambridge University Press, Chap. 5, 2010.

[22] Burt, Peter J., and Edward H. Adelson. "The Laplacian pyramid as a compact image code." Communications, IEEE Transactions on 31, no. 4 (1983): 532-540.

[23] Minh N. Do and Martin Vetterli, The Finite Ridgelet Transform for Image Representation, IEEE Transactions on Image Processing, 12(1), January 2003, 16-28.

[24] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," IEEE Signal Process. Mag., vol. 18, pp. 36–58, Sep. 2001.

[25] S. R. Deans, The Radon transform and some of its applications, Wiley, 1983.

[26] D. L. Donoho, M.Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," IEEE Trans. Inf. Theory, vol. 44, no. 6, pp. 2435–2476, Oct. 1998.

[27] Candes, Emmanuel Jean. "Ridgelets: theory and applications." PhD diss., Stanford University, 1998.

[28] D. L. Donoho, "Orthonormal ridgelets and linear singularities," SIAM Journal on Mathematical Analysis, 31, 2000, Citeseer, pp. 1062-1099.

[29] E. J. Candes, D. L. Donoho, and Stanford University Department of Statistics., Curvelets: A surprisingly effective nonadaptive representation for objects with edges, Citeseer, 2000.

[30] C. Anteneodo and C. Tsallis, "Multiplicative noise: a mechanism leading to nonextensive statistical mechanics", arXiv:cond-mat/0205314 v2, 2003; Journal of Mathematical Physics, 44, 5194–5203, 2003.

[31] Julesz, B., "Visual Pattern Discrimination," Information Theory, IRE Transactions on , vol.8, no.2, pp.84,92, February 1962.

[32] I. Fogel and D. Sagi, "Gabor filters as texture discriminators," Biol. Cybern. 61, 103-113 (1989).

[33] Rosenholtz R. Texture perception. In: Wagemans J, editor. Oxford Handbook of Perceptual Organization (in press). Oxford, U.K: Oxford University Press; 2014.

[34] F. Murtagh and J-C. Starck, "Wavelet and Curvelet Moments for Image Classification: Application to Aggregate Mixture Grading," Pattern Recognition Letters, vol. 29, pp. 1557-1564, 2008.

[35] Mosquera-Lopez, Clara, Sos Agaian, and Alejandro Velez-Hoyos. "The development of a multi-stage learning scheme using new tissue descriptors for automatic grading of prostatic carcinoma." In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on, pp. 3586-3590. IEEE, 2014.

[36] D. F. Gleason, and G. T. Mellinger, "The Veterans Administration Cooperative Urological Research Group: Prediction of Prognosis for Prostatic Adenocarcinoma by Combined Histological Grading and Clinical Staging," J Urol, v.111, pp. 58-64, 1974.

[37] Luthringer, D. J., and Gross, M.,"Gleason Grade Migration: Changes in Prostate Cancer Grade in the Contemporary Era," PCRI Insights, vol. 9, pp. 2-3, August 2006.

[38] J.I. Epstein, "An update of the Gleason grading system," J Urology, v. 183, pp. 433-440, 2010.

[39] Pierorazio PM, Walsh PC, Partin AW, and Epstein JI., "Prognostic Gleason grade grouping: data based on the modified Gleason scoring system," BJU International, 2013.

[40] Zhang, M. M. Islam, G. Lu and I. J. Sumana, "Rotation Invariant Curvelet Features for Region Based Image Retrieval," Intern. J. Comput. Vis., vo. 98, pp. 187-201, 2012.

[41] Zand, Mohsen, et al. "Texture classification and discrimination for region-based image retrieval." Journal of Visual Communication and Image Representation 26, pp. 305-316. 2015.

[42] F. Gomez and E. Romero, "Rotation Invariant Texture Classification using a Curvelet based Descriptor," Pattern Recog. Lett, vol. 32, pp. 2178-2186, 2011.

[43] S. Arivazhagan and T. G. S. Kumar, "Texture Classification using Curvelet Transform," Intern. J. Wavelets, Multiresolution & Inform Processing, vol. 5, pp. 451-464, 2007.

[44] S. Arivazhagan, L. Ganesan and T. G. S. Kumar, "Texture Classification using Curvelet Statristical and Co-occurrence Features," Proc. IEEE ICPR'06, pp. 938-941, 2006.

[45] L. Shen and Q. Yin, "Texture Classification using Curvelet Transform," Proc. ISIP'09, China, pp. 319-324, 2009.

[46] M. Unser and M. Eden, "Multiresolution Texture Extraction and Selection for Texture Segmentation," IEEE Trans. PAMI, vol. 11, pp. 717-728, 1989.

[47] M. Unser, "Texture Classification and Segmentation using Wavelet Frames," IEEE Trans. IP, vol. 4, pp. 1549-1560, 1995.

[48] Siegel, R. L., Miller, K. D. and Jemal, A. (2017), Cancer Statistics, 2017. CA: A Cancer Journal for Clinicians, 67: 7–30. doi:10.3322/caac.21387.

[49] D. F. Gleason, and G. T. Mellinger, "The Veterans Administration Cooperative Urological Research Group: Prediction of Prognosis for Prostatic Adenocarcinoma by Combined Histological Grading and Clinical Staging," J Urol, v.111, pp. 58-64, 1974.

[50] Luthinger, D. J., and Gross, M., "Gleason Grade Migration: Changes in Prostate Cancer Grade in the Contemporary Era," PCRI Insights, vol. 9, pp. 2-3, August 2006.

[51] J.I. Epstein, "An update of the Gleason grading system," J Urology, v. 183, pp. 433-440, 2010.

[52] Pierorazio PM, Walsh PC, Partin AW, and Epstein JI., "Prognostic Gleason grade grouping: data based on the modified Gleason scoring system," BJU International, 2013.

[53] K. Jafari-Khouzani, and H. Soltanian-Zadeh, "Multiwavelet Grading of Pathological Images of Prostate," IEEE Trans. on Bio. Eng., v.50, pp. 697-704, 2003.

[54] A. Tabesh, M. Teverovskiy, H. Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, "Multifeature Prostate Cancer Diagnosis and Gleason Grading of Histological Images," IEEE Trans. on Medical Imaging, v. 26, pp. 1366-1378, 2007.

[55]   S. Doyle, M. Feldman, J. Tomaszewski and A. Madabhushi, "A boosted Bayesian multi-resolution classifier for prostate cancer detection from digitized neeedle biopsies," IEEE Trans. On Biomed. Eng., v. 57, n. 5, pp. 1205-1216, 2012.

[56]   Hong-Jun Yoon, Ching-Chung Li, C. Christudass, R. Veltri, J. I. Epstein, and Zhen Zhang, "Cardinal Multiridgelet-based Prostate Cancer Histological Image Classification for Gleason Grading," PROC. IEEE BIBM, pp. 315-320, 2011.

[57]   Nguyen, Kien, Anil K. Jain, and Bikash Sabata, "Prostate cancer detection: Fusion of cytological and textural features," Journal of pathology informatics 2, 2011.

[58]   K. Nguyen, B. Sabata and A. K. Jain, "Prostate Cancer Grading: Gland Segmentation and structural Features," Pattern Recognition Letter, v.33, Issue 7, pp.951-961, 2012.

[59]   S. Ali, R. Veltri, J. Epstein, C. Christudass and A. Madabhushi, "Adaptive Enegy Selective Active Contour with Shape Priors for Nuclear Segmentation and Gleason Grading of Prostate," Proc. MICCAI, pp. 661-669, 2011.

[60]   Veltri, Robert W., Christhunesa S. Christudass, and Sumit Isharwal, "Nuclear morphometry, nucleomics and prostate cancer progression," Asian journal of andrology 14.3, pp. 375-84, 2012.

[61]   A. Bezerianos, S. Tong and N. Thakor, "Time-dependent entropy estimation of EEG rhythm changes following brain ischemia", Annals of Biomedical Engineering, 31, 221–322, 2003.

[62]   M.K. Tsatsanis and G.B. Giannakis, "Object and texture classification using higher order statistics", IEEE Transactions on Pattern Analysis and Machine Intelligence, 14, 733–750, 1992.

[63]   K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video," IEEE Trans. Circuits Sys. Video Tech., Vol. 5, pp.52-56, 1995.

[64]   N. Ponomarenko, V. Lukin, S. Abramov, K. Egiazarian, J. Astola, "Blind evaluation of additive noise variance in textured images by nonlinear processing of block DCT coefficients," Image Processing: Algorithms and Systems II: proc. of Internat. Conf., USA, 2003, SPIE Vol. 5014, pp. 178-189.

[65]   P. Khurd, C. Bahlmann, P. Maday, A. Kamen, S. Gibbs-Strauss, E. M. Genega, and J. V. Frangioni, "Computer-aided Gleason grading of prostate cancer histopathological images using texton forests," 2010, IEEE, pp. 636-639.

[66]   http://urology.jhu.edu/research/veltri/today.php

[67]   Smith, Steven W. "The scientist and engineer's guide to digital signal processing.", 1997.

[68]   Coetzee, Frans M. "Correcting the Kullback–Leibler distance for feature selection." Pattern Recognition Letters 26.11 2005: 1675-1683.

[69]   Candè, Emmanuel J., and Michael B. Wakin. "An introduction to compressive sampling." Signal Processing Magazine, IEEE 25, no. 2 (2008): 21-30.

[70]   Epstein, Jonathan I. "A new contemporary prostate cancer grading system." Pathology international (2015).

[71]   Epstein, Jonathan I., William C. Allsbrook Jr, Mahul B. Amin, Lars L. Egevad, and ISUP Grading Committee. "The 2005 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma." The American journal of surgical pathology 29, no. 9 (2005): 1228-1242.

[72]   Sperling, MD, D. "Revisions of the Gleason grading system make it more accurate". Sperling Prostate Center. Retrieved 22 April 2016.

[73]   A. Tabesh, M. Teverovskiy, H. Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, "Multifeature Prostate Cancer Diagnosis and Gleason Grading of Histological Images," IEEE Trans. on Medical Imaging, v. 26, pp. 1366-1378, 2007.

[74]   Diaconescu, S., D. Diaconescu, and S. Toma. "Nucleolar morphometry in prostate cancer." Bull Transilvania Univ Brasov 3 (2010): 23-6.

[75]   Zink, Daniele, Andrew H. Fischer, and Jeffrey A. Nickerson. "Nuclear structure in cancer cells." Nature reviews cancer 4, no. 9 (2004): 677-687

[76]   Veltri, Robert W., et al. "Ability to predict metastasis based on pathology findings and alterations in nuclear structure of normal-appearing and cancer peripheral zone epithelium in the prostate." Clinical Cancer Research 10.10 (2004): 3465-3473.

[77]   Laws, Kenneth I. Textured image segmentation. No. USCIPI-940. University of Southern California Los Angeles Image Processing INST, 1980.

[78]   Sebai, Houria, and Assia Kourgli. "A comparative study of feature vectors derived from wavelets applied to high resolution satellite images retrieval." In 2014 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1-6. IEEE, 2014.

[79]   Materka, Andrzej, and Michal Strzelecki. "Texture analysis methods–a review." Technical university of lodz, institute of electronics, COST B11 report, Brussels (1998): 9-11.

[80]   Adelson, Edward H., Charles H. Anderson, James R. Bergen, Peter J. Burt, and Joan M. Ogden. "Pyramid methods in image processing." RCA engineer 29, no. 6 (1984): 33-41.

[81]   Y. Meyer. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. Séminaire Bourbaki, 662, 1986.

[82]   Daubechies, Ingrid. Ten lectures on wavelets. Vol. 61. Philadelphia: Society for industrial and applied mathematics, 1992.

[83]   Alecu, Alin, Adrian Munteanu, Aleksandra Pizurica, Wilfried Philips, Jan Cornelis, and Peter Schelkens. "Information-theoretic analysis of dependencies between curvelet coefficients." In Image Processing, 2006 IEEE International Conference on, pp. 1617-1620. IEEE, 2006.

[84]   Andrew G. Bruce, Hong-Ye Gao. MathSoft, Inc. (USA). Proc. SPIE 2569, Wavelet Applications in Signal and Image Processing III, 270 (September 1, 1995).

[85]   Gao, Hong-Ye. "Wavelet shrinkage denoising using the non-negative garrote." Journal of Computational and Graphical Statistics 7, no. 4 (1998): 469-488.

[86]   Breiman, Leo. "Better subset regression using the nonnegative garrote." Technometrics 37, no. 4 (1995): 373-384.

[87]   Yi, Ting-Hua, Hong-Nan Li, and Xiao-Yan Zhao. "Noise smoothing for structural vibration test signals using an improved wavelet thresholding technique." Sensors 12, no. 8 (2012): 11205-11220.

[88]   Sendur, Levent, and Ivan W. Selesnick. "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency." IEEE Transactions on signal processing 50, no. 11 (2002): 2744-2756.

[89]   Donoho, David L., and Jain M. Johnstone. "Ideal spatial adaptation by wavelet shrinkage." Biometrika 81, no. 3 (1994): 425-455.

[90]   Cai, T. Tony, and Bernard W. Silverman. "Incorporating information on neighbouring coefficients into wavelet estimation." Sankhyā: The Indian Journal of Statistics, Series B (2001): 127-148.

[91]   Chang, S. Grace, Bin Yu, and Martin Vetterli. "Adaptive wavelet thresholding for image denoising and compression." IEEE transactions on image processing 9, no. 9 (2000): 1532-1546.

[92]   Donoho, David L., and Iain M. Johnstone. "Adapting to unknown smoothness via wavelet shrinkage." Journal of the American statistical association 90, no. 432 (1995): 1200-1224.

[93]   C. M. Stein, "Estimation of the mean of a multivariate normal distribution," Ann. Statist., vol. 9, no. 6, pp. 1135–1151, 1981.

[94]   Abramovich, Felix, Theofanis Sapatinas, and Bernard W. Silverman. "Wavelet thresholding via a Bayesian approach." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 60, no. 4 (1998): 725-749.

[95]   Canny, John. "A computational approach to edge detection." IEEE Transactions on pattern analysis and machine intelligence 6 (1986): 679-698.

[96]   Vacchetti, Luca, Vincent Lepetit, and Pascal Fua. "Combining edge and texture information for real-time accurate 3d camera tracking." In *Mixed and Augmented Reality, 2004. ISMAR 2004. Third IEEE and ACM International Symposium on*, pp. 48-56. IEEE, 2004.

[97]   Ching-Chung Li and Wen-Chyi, "On Curvelet-Based Texture Features for Pattern Classification," in *Handbook of Pattern Recognition and Computer Vision*, 5th ed., Eds., C. H. Chen, Chap. 1.7., World Scientific.

[98]   Veltri, Robert W., Sahirzeeshan Ali, Wen-Chyi Lin, Guangjing Zhu, Jonathan I. Epstein, Ching-Chung Li, and Anant Madabhushi. "Cancer histologic and cell nucleus architecture differentiate prostate cancer Gleason patterns 3 from 4."Cancer Research 75, no. 15 Supplement (2015): 4349-4349.

[99]   P. Khurd, L. Grady, A. Kamen, S. Gibbs-Strauss, E. M. Genega and J. V.Frangioni, "Network cycle features: application to computer-aided Gleason grading of prostate cancer histopathological images," Proc. IEEE ISBI, 2011 pp.1632-1626.

[100] Fehr, Duc, Harini Veeraraghavan, Andreas Wibmer, Tatsuo Gondo, Kazuhiro Matsumoto, Herbert Alberto Vargas, Evis Sala, Hedvig Hricak, and Joseph O. Deasy. "Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images." Proceedings of the National Academy of Sciences 112, no. 46, 2015, pp.6265-6273.

[101] Ji, Hui, Xiong Yang, Haibin Ling, and Yong Xu. "Wavelet domain multifractal analysis for static and dynamic texture classification." IEEE Transactions on Image Processing 22, no. 1 (2013): 286-299.

[102] Chan, Tony F., and Luminita A. Vese. "Active contours without edges." IEEE Transactions on image processing 10, no. 2 (2001): 266-277.

[103] Zong, Yang, Jiaoti Huang, Devipriya Sankarasharma, Teppei Morikawa, Masashi Fukayama, Jonathan I. Epstein, Kiran K. Chada, and Owen N. Witte. "Stromal epigenetic dysregulation is sufficient to initiate mouse prostate cancer via paracrine Wnt signaling." Proceedings of the National Academy of Sciences 109, no. 50 (2012): E3395-E3404.

[104] Lin, Wen-Chyi, Ching-Chung Li, Jonathan I. Epstein, and Robert W. Veltri. "Curvelet-based texture classification of critical Gleason patterns of prostate histological images." In Computational Advances in Bio and Medical Sciences (ICCABS), 2016 IEEE 6th International Conference on, pp. 1-6. IEEE, 2016. DOI: 10.1109/ICCABS.2016.7802768.

[105] Schölkopf, Bernhard, and Alexander J. Smola. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press, 2002.

[106] Brochu, Eric, Vlad M. Cora, and Nando De Freitas. "A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning." arXiv preprint arXiv:1012.2599 (2010).

[107]   Bergstra, James, and Yoshua Bengio. "Random search for hyper-parameter optimization." Journal of Machine Learning Research 13, no. Feb (2012): 281-305.

[108]   https://www.mathworks.com/help/pdf_doc/stats/stats.pdf

[109]   Azar, Jimmy C., Martin Simonsson, Ewert Bengtsson, and Anders Hast. "Automated classification of glandular tissue by statistical proximity sampling." Journal of Biomedical Imaging 2015 (2015): 3.

[110]   K. Nguyen, A. Sarkar and A. K. Jain, "Prostate Cancer Grading: Use of Graph Cut and Spatial Arrangement of Nuclei," IEEE Trans. Medical Imaging, vol. 33, no. 12, pp. 2254-2270, Dec. 2014.

[111]   Naik, Shivang, et al. "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology." Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on. IEEE, 2008, pp. 284–287.