JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION
Vol. 5, No. 1, March, pp. 41-51, 1994

Robust Spatial Autoregressive Modeling for Hardwood Log Inspection

DONGPING ZHU[1] AND A. A. (LOUIS) BEEX

Bradley Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0111

Received December 19, 1991; accepted January 27, 1993

We explore the application of a stochastic texture modeling method toward a machine vision system for log inspection in the forest products industry. This machine vision system uses computerized tomography (CT) imaging to locate and identify internal defects in hardwood logs. The application of CT to such industrial vision problems requires efficient and robust image analysis methods. This paper addresses one particular aspect of the problem of creating such a computer vision system, namely, the use of image texture modeling for wood defect recognition. In particular, we contribute the first application of spatial autoregressive (SAR) modeling to wood-grain texture analysis of CT images of hardwood logs. Thereto a circularly shifted correlation approach is developed to discriminate the circular texture patterns on the cross-sectional CT images of logs. A robust algorithm for parameter estimation is applied to obtain model parameters associated with individual defects occurring inside a log. Based on the estimated model features, a simple minimum distance correlation-classifier is constructed which classifies an unknown defect into one of the prototypical defects. Experimental results from the proposed method, applied to CT images from different red oak wood, are given and show the efficacy of our approach.

I. INTRODUCTION

The research reported in this work on a computer vision system for automatically processing hardwood logs is to help the hardwood sawmill industry automate, reduce costs, and increase product volume and value recovery. Log inspection is currently performed by human experts. Their performance is, however, quite limited due to a number of factors. More quantitative approaches are called for to automate grading and board cutting planning of logs. A key to the success of such work is the ability to automate the detection of defects inside logs. After the defects are located, properly sized, and identified, board outlines must be determined as well. Subsequently computer programs can grade the logs and determine possibilities for cutting boards.

[1]Present address: InVision Technologies, Inc., Foster City, CA.

Technology has developed to the point that wood defects can indeed be "seen" in the interior of logs using non-destructive imaging techniques. More recently, log scanning using sophisticated methods, such as NMR (nuclear magnetic resonance) and CT (computerized tomography) imaging, has been studied in laboratory experiments. The ability to differentiate some internal defects has been demonstrated in part [1-5].

There are two important characteristics in automatically processing logs. First, log defects must be represented and analyzed as three-dimensional data. Individual images are captured and stored as two-dimensional slices and most of the image analysis task takes place with these slices. Three-dimensional images are then created by reconstructing a sequence of two-dimensional images or slices of a log. Second, the computer vision system must be able to detect features of a log that are not visible to the human observer. This invisibility aspect makes research in this area difficult and time consuming because logs must eventually be cut to verify the existence, location, and size of the defects that one has anticipated or missed with the vision system.

A nondestructive way of inferring the internal structure of a three-dimensional object, such as a log, is to use computerized tomography (CT) to calculate the attenuation of X-ray transmission for each small volume of the object. CT imaging of a hardwood log along its length produces a stack of cross-sectional slices representing the three-dimensional structure of the log. The image gray-level value at each pixel is called the CT number. Typical CT numbers range from – 1024 to 1024 (11-bit data), where water has a value of O. Since X-ray attenuation depends on the density of the material, CT numbers represent density measurements of the wood structure in a log. The proposed computer vision system is designed to be independent of wood species, and consists of two basic modules: the low level module performs such tasks as image filtering, segmentation, region detection and merging [5], and the high level module conducts defect recognition. The final output of the proposed vision system, for a particular log, would be a data base that describes the distribution of defects inside that log. This

1047-3203/94 $6.00

database is then forwarded to a numerical control unit (NCU) where the actual cutting takes place.

Although industrial interest in log inspection using CT imaging technology has been increasing, there are few existing practical systems known to the authors. Available literature on analyzing CT images of, specifically, hardwood logs is sparse.

Related information pertains to the detection of defects in softwood logs [3]. CT scans of a log were taken with a Siemens medical scanner. To identify different regions on each CT slice, three features were derived and used in the recognition process: (1) pixel value, (2) object shape, and (3) growth-ring texture patterns. In its decision-making process, the system discerns knots by their high pixel value and 2-D shape, good wood by the circular patterns of the growth-rings, rot by its high pixel value and growth-ring patterns, and background (air) by its low pixel value. The system can successfully identify knots, rot holes, and good wood in CT scans of softwood sawlogs. As a texture feature, the uniformity of the growth-ring texture is used. The first step in the test of growth-ring uniformity is to find individual edge elements with a standard edge detection algorithm. For each pixel that is on an edge, the directions to the neighboring pixels on the same edge are averaged to yield a direction for the edge at the current pixel. The edge directions at all pixels within a small local area are then compared. The higher the percentage having the same direction, the higher the growth-ring uniformity is taken to be.

The spatial autoregressive (SAR) model-based texture analysis method has been applied previously to the grading of *surface defects* in planed hardwood boards [4]. The image data of the wood boards were acquired using an industrial camera, rather than a CT scanner. Although similar texture models were used in that study, the nature of the image data and the formation of the wood textures are both significantly different from those in the present stud y. The wood grain textures were imaged in the longitudinal plane of the logs [4], rather than in the cross-sectional plane. The texture features utilized for pattern classification were the SAR model parameter estimates and the image sample mean on 64 × 64 subimages [4], rather than the SAR model parameters, the estimated error variance, and the image sample mean and variance on 3 × 3 subimages. For the classification of board defects three features were selected from the feature vector, using the scatter matrix algorithm. The selected features were linear combinations of the members of the feature vectors. However, the a priori probability of each class of textures is needed to compute the within- and between-class scatter matrices for the feature selection. Equal probability of occurrence for each texture class was assumed. This subjective assumption is not necessarily valid in practice, so that a more accurate measure of the a priori class probability must be developed. In contrast to the tree classifier [4], we propose a pair-wise pattern classification method which does not require the a priori probability of occurrence of the texture classes under examination. Therefore, we see the work reported here, as applied to CT images of hardwood logs, as an extension of the work on planed boards [4].

Section II gives economic and practical considerations for the overall log inspection vision system. Texture modeling in general is treated and contrasted with SAR modeling in particular in SectionIII. We explore the application of Spatial Auto-Regressive modeling to the texture analysis of CT images of hardwood logs, in order to locate and identify internal defects in those hardwood logs. Thereto we develop, in particular, a circularly shifted correlation approach to discriminate the circular texture patterns on these cross-sectional CT images. Details on SAR modeling and its robust identification are given in Section IV. Section V covers the proposed SAR model-based correlation-classification procedure. A simple minimum distance classifier is constructed, based on the estimated model features, which classifies an unknown defect into one of the prototypical defects. The experimental results, with CT images from different red oak wood, some analysis thereof, and a resulting improved classifier are given in Section VI. The results with the improved classifier show the efficacy of the proposed approach. Section VII provides concluding remarks.

II. FEASIBILITY, PRACTICALITY, AND COST EFFECTIVENESS

The important components of a vision system for log inspection include a CT scanner, special purpose data acquisition and conversion hardware, and image analysis software. The CT scanner is the core component of the overall system, which determines the image quality and dominates the speed performance of the system.

Important system issues for the practical implementation of the vision system include imaging quality, scanning rate, efficient soft ware implementation, and portability of the scanning system. *Imaging quality* of a CT scanner depends on a number of factors, such as a stable high voltage power supply, a durable X-ray tube, good mechanical alignment, and an efficient reconstruction algorithm. Among these, the power supply and reconstruction algorithm are the most important factors. Since X-ray attenuation (which is reflected in image contrast) increases, and image signal-to-noise ratio (SNR) decreases, as the power supply increases, a trade-off between image contrast and SNR is achieved when the power voltage is in the range from 80 to 200 kV. Another concern with the power supply is the fact that an unstable power supply tends to produce reconstruction artifacts, such as streaks.

The third and fourth generation CT scanners have a *scanning rate* of one slice per second. Since a rate of 0.5 log per min is not uncommon in many sawmills, little time is available in which to acquire and analyze an extremely large amount of CT image data. Recently, more rapid scanning is made possible because of a new scanning electron beam CT scanner [6]. In its ultrafast mode, the fifth generation scanner acquires a pair of slices at the rate of 34 images per s. This improvement of scanning speed is significant and will revolutionize the log scanning industry.

Efficiency of software implementation results when the reconstruction algorithms, which are typically Fourier transform-based, are implemented on high speed, special-purpose array processors. In addition, some of the image analysis software can be implemented in real-time using digital hardware, e.g., a histogram-based segmentation has been efficiently implemented using a DMA (direct memory access) technique on an IBM PS2 AIX system in the Spatial Data Analysis Laboratory at Virginia Tech. Other image analysis computation, such as the robust identification, can also be accomplished fast via programming of one or several of the high-speed DSP (digital signal processing) chips currently available.

The last issue for implementation of CT-based scanning systems is the *portability* of the system. For normal log scanning, the diameter of the scanner gantry ranges from 30 to 60 cm. A conveyor is needed to move logs up and down the scanner at high speed. The whole scanning system can be packed in a mobile trainer so that it can be readily moved from one log yard to another. One of the examples is the portable CTX500 system, deployed now at the San Francisco International Airport by the Federal Aviation Administration. Special industrial CT scanners have been built at a cost of about $4000 a piece at Chrisholm Institute of Technology in Australia.

The knowledge of internal log defects is critical in deciding whether to veneer or to saw up a log, and how to position a log so that the boards sawn from it will have as much clear-face as possible. This will in turn produce lumber of the highest value. Tree-length round-wood may need to be cut prior to further processing into lumber or veneer. With internal defect information, it becomes possible to cut round-wood so that defects are removed from the log or isolated at either end. This leaves larger areas of valuable clear-wood in the log, and gives it higher value. At least 10 to 20% improvement in lumber value would result if an optimal cutting strategy could be formulated [7–9]. At the same time certain logs or lengths of a log can be identified as veneer quality; utilization as veneer can increase log value by a factor of 10.

Log costs have increased faster than lumber prices in the last 30 years; in fact, from 20 to 80% of the total production cost. Lumber processing facilities in the Eastern United States alone produce more than 10 billion board feet of hardwood lumber annually. Therefore, improving lumber processing by inspecting saw-logs would not only raise mill profits, but also enhance the utilization of the hardwood resource. Even a small amount of value recovery increase per log, achieved through log scanning, would turn into great economic margin when scanning millions of logs in sawmills.

Taking into account the image analysis software development costs and other manufacturing costs, a CT-based system for automated log inspection appears feasible, practical, as well as cost-effective.

III. TEXTURE MODELING

Image texture is an important feature used in analyzing objects or regions of interest in image analysis and computer vision applications. Texture characteristics play a major role in image classification and analysis. In classification, texture features can be used to discriminate and label regions of an image, such as crop identification in an aerial photograph, and medical diagnosis of an X-ray photograph. Texture features can also be used in scene segmentation and identification in an image understanding or computer vision system, for example, in robot vision and industrial inspection. Therefore, the determination of proper texture features is the key to these applications,

Some of the currently used choices in texture features include [10]: (1) features derived from second-order gray-level statistics such as the gray-level co-occurrence matrices, (2) gray-level run-length statistics, (3) gray-level difference statistics, (4) moments derived from decorrelation methods [11], and (5) rings and wedges derived from the Fourier power spectrum [12].

Texture classification using *gray-level co-occurrence matrices* [13] is based on the premise that human perception rely on the second-order statistics as reflected in these matrices. However, such matrices are not used directly. Instead, derived measures, such as energy and entropy, etc., are computed and used as discriminating features. In the *gray-level run-length method* the numbers of occurrences of a specific value of a particular length are used to build a matrix. Different orientations of the runs produce different matrices. Features, such as short and long run emphasis, etc., are computed and used for discriminating textures. Due to the large amount of matrix computation, the method is especially suitable for classifying texture patterns with a limited number of orientations. In the *gray-level difference method* one computes the occurrence of difference of two gray-levels separated by a specific distance. It so generates a set of probability distributions from which features like contrast, second moment, etc., are derived and used for classification. Therefore, information about the gray-level is not used directly. In the *decorrelation method of feature*

extraction the texture is modeled as a separable causal Markov field. From the point of view of image modeling, in the limit, the decorrelation features are extracted from the moments of residuals obtained by fitting the 2-D model to the given texture. The loss of information during this process of dimensionality reduction may be significant since the residuals themselves contain only partial information regarding the original data. In practice, however, the Gaussian assumption for the feature vectors with known mean and variance matrices is not always true, which limits its application to industrial inspection. *Fourier domain features extracted from the power spectrum* can also be used to discriminate textures. The power spectrum can be computed from the periodogram, and features obtained from it, such as rings and wedges. The power spectrum can be expressed in polar coordinates [14], which for directional textures yields peaks as a function of the angle.

In this work, the SAR model [15] is adopted to characterize wood-grain textures stochastically, since there exists a high interpixel correlation as well as a strong growth directionality in the wood-grain patterns. Different wood defects will exhibit different texture patterns, which can be systematically modeled by a parametric texture model with different parameters. The model parameters can then be used to discriminate different wood defects. Another advantage of using the parametric SAR model is that it can be readily adopted, in the texture pattern recognition step, to detect the directionality of the wood grain pattern. While SAR models have been used to analyze textures [4, 15, 16], no study has dealt with the particular texture that exhibits circular patterns. To cope with this form of directionality, we develop a pattern recognition method based on the concept of circularly shifted correlation between two feature vectors.

In comparison with the methods mentioned in the beginning of this section then, the texture classification approach used in this paper is characterized by: (1) a non-Gaussian assumption for the noise sequence, (2) reduction of the data to a reasonable amount, (3) utilization of the directional information of textures, and (4) a simple design for the pattern classifier.

Finally, most pattern classification techniques are nonparametric clustering methods [17] that maximize (minimize) an objective function. A fuzzy set approach to clustering is based on the premise that a sample does not necessarily belong to one particular class, but rather exhibits class membership values [18], indicating the confidence with which it is associated with any of the possible classes, Fuzzy clustering algorithms do not assume any model for the observations under examination and as such they may be considered nonparametric techniques. Without parameterization of the data set by a compact set of parameters, the minimization is usually performed with respect to all the class membership values.

In our SAR parameterization the uncertainty of the observations is characterized by an inexact underlying model, and methods of robust statistics are used to make the classification algorithm robust. In that way, the method we present fills in the gap between nonparametric methods and fully parametric techniques (which assume an exact parametric form of the underlying distribution of the sample). The maximum likelihood method is a well-established technique to estimate unknown parameters and establish the sample label. That technique is extended here by using the weighted M-estimator [19], which makes the classification procedure robust and results in a reliable method.

IV. SAR MODELING AND IDENTIFICATION

The cross-sectional CT images of a log contain the fine structural details of annual rings which tend to hinder the later image processing steps of the vision system. An efficient three-dimensional adaptive filtering method is described in [5] that can eliminate the unwanted ring structures while preserving other important image details, e.g., the presence of small checks. The filtered images are then segmented on an image-by-image basis, producing a number of regions each representing a potential defect such as splits, knots, barks, stains, and decays [5]. For each defect region of unknown class on a CT log image that has been detected by the segmentation module, a Gaussian spatial autoregressive (SAR) model is defined on a neighborhood system to model the spatial directional dependence of wood-grain, i.e., each pixel value is modeled as a weighted sum of the values of its neighboring pixels plus a Gaussian white noise process. For each homogeneous region, the SAR model parameters are estimated using a generalized M-estimation approach. This set of SAR parameters as well as the gray-level mean and variance are then employed as a feature vector in a supervised texture discrimination scheme.

The subset of gray-values in an image region Ω is represented by $\{y_s | s = (i, j) \in \Omega\}$. Currently a square window is chosen, for simplicity, so that $\Omega = \{(i, j) | 1 \le i \le M,\ 1 \le j \le M\}$. The method is directly applicable to other—even arbitrarily shaped—windows, however, which is indicated in the identification equations later in this section by letting the argument of the summations range over $s \in \Omega$.

An independent Gaussian random field [16] can be represented by

$$y_s = f + w_s, \tag{1}$$

where f is a constant signifying the image field mean value, w_s is a two-dimensional zero-mean Gaussian white noise process with unknown variance σ_w^2. Under this model assumption, f and σ_w^2 are the feature values that

represent the textural image field. Although this model has been used to model certain simple textures, it could not fully characterize the wood-grain textures under study.

In the following, a second type of random field is used as the model for defect recognition. It is assumed that the subset $\{y_s\}$ satisfies the so called Gaussian causal or spatial autoregressive (CAR or SAR) model in a toroidal lattice system [4],

$$y_s = \boldsymbol{\theta}^{\mathrm{T}}\mathbf{y}_s + \sigma w_s, \tag{2}$$

where $\boldsymbol{\theta}$ is the SAR parameter vector consisting of p unknown coefficients θ_i, $i = 1, 2, \ldots, p$; $\mathbf{y}_s$ is the neighborhood vector containing the lexicographically ordered image data from the neighborhood of s, for a particular definition of neighborhood system, and w_s is a zero-mean, unit-variance white noise process. The model parameter vector $\boldsymbol{\theta}$, the model input variance σ^2, the image gray-level mean μ_y, and the image gray-level variance σ_y^2 are to be estimated from the CT image, and together they comprise the feature vector for each individual wood texture.

To account for the unrealistic Gaussian assumption for the image data in a potential region, a contamination model [4, 19] is introduced to take into consideration the effects of *outliers*, the data points that are (statistically) far away from their assumed cluster centers. In this approach, the observed image samples z_s satisfy

$$z_s = y_s + u_s, \tag{3}$$

where the noise process u_s is independent of y_s and identically distributed with a mixture density H given by

$$H = (1 - \alpha)F + \alpha G \tag{4}$$

consisting of two components: a degenerate central component F and a contamination component G. For instance, F can be a Gaussian distribution with zero mean and variance σ_F^2, and G an arbitrary distribution characterizing the outliers in the image data, such as the zero mean Laplacian distribution. The iterative algorithm described below, however, does not depend on the explicit forms of F and G. In the above, α is a small fraction of contamination (typically <0.2), defined as the ratio of the number of outliers in the data to the total number of data. The model expressed (4) states that the majority of the image data (a fraction (1-α)) is distributed according to F, and a minority of the image data (a fraction α) is distributed according to G.

In the robust estimation method proposed in [19], the texture model parameters $\{\boldsymbol{\theta}, \sigma\}$ in (2), are determined from the observed image z_s by minimizing the index $J(\boldsymbol{\theta}, \sigma)$,

$$J(\boldsymbol{\theta}, \sigma) = \sum_{s\in\Omega} W1_s \left[\rho\left(\frac{r_s}{\sigma}\right) + q\right]\sigma, \tag{5a}$$

where $W1_s$ is a weighting function, ρ is the Hampel robustifying function (specified later), q is an optimization constant, and r_s is the residual defined as

$$r_s = z_s - \boldsymbol{\theta}^{\mathrm{T}}\mathbf{z}_s, \tag{5b}$$

where $\mathbf{z}_s$ is the neighborhood vector derived from $\mathbf{y}_s$, by replacing its elements with observed image values. The objective here is to minimize $J(\boldsymbol{\theta}, \sigma)$ in (5) with respect to the parameter vector $\boldsymbol{\theta}$ and the standard deviation σ. From using a stochastic approximation in the steepest descent gradient method [4, 20], the following iterative algorithm results for the simultaneous estimation of the model parameter vector $\boldsymbol{\theta}_k$ and the model input variance σ_k^2, also referred to as the estimated error variance, at each iteration index $k = 1, 2, 3, \ldots$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \gamma\left[\sum_{s\in\Omega} \mathbf{z}_s W1_s \psi'\left(\frac{r_{s,k}}{\sigma_k}\right)\mathbf{z}_s^T\right]^{-1} \cdot \left[\sum_{s\in\Omega} \mathbf{z}_s W1_s W2\left(\frac{r_{s,k}}{\sigma_k}\right) r_{s,k}\right] \tag{6a}$$

$$\sigma_{k+1}^2 = \sum_{s\in\Omega} W1_s \chi\left(\frac{r_{s,k}}{\sigma_k}\right)\sigma_k^2 \Big/ \sum_{s\in\Omega} \beta W1_s \tag{6b}$$

with γ being a convergence step size satisfying $0 < \gamma < 2$, and the weighting functions $W1_s$ and $W2_s$ as

$$W1_s = \frac{\psi(\sqrt{\mathbf{z}_s^{\mathrm{T}}\mathbf{z}_s/\sigma_y^2 p})}{\sqrt{\mathbf{z}_s^{\mathrm{T}}\mathbf{z}_s/\sigma_y^2 p}} \tag{7}$$

$$W2_x = \frac{\psi(x)}{x}, \tag{8}$$

where p is the number of elements in the parameter vector $\boldsymbol{\theta}$. In addition

$$\chi(x) = x\psi(x) - \rho(x) \tag{9}$$

$$\beta = E_\Phi\{\psi^2(r_s)\}, \tag{10}$$

where E_Φ denotes the expectation over the standard-normal distribution $N(0, 1)$. In the above equations, σ_y^2 is the gray-level value variance calculated over the image region Ω, and the odd function $\psi(x)$ is

$$\psi(x) = \begin{cases} x & |x| \le a \\ a\ \mathrm{sign}(x) & a < |x| \le b \\ a\ \mathrm{sign}(x)\dfrac{c - |x|}{c - b} & b < |x| \le c \\ 0 & c < |x|, \end{cases} \tag{11}$$

which is the first derivative of the, even, Hampel robustifying function $\rho(x)$

$$\rho(x) = \begin{cases} \dfrac{x^2}{2} & |x| \le a \\ a|x| - \dfrac{a^2}{2} & a < |x| \le b \\ a\dfrac{c|x| - (x^2/2)}{c-b} - \dfrac{a^2}{2} - \dfrac{ab^2}{2(c-b)} & b < |x| \le c \\ \dfrac{a}{2}(c + b - a) & c < |x|. \end{cases} \tag{12}$$

The parameters in $\psi(x)$ and $\rho(x)$, used for the experiment section, are $a = 2.0$, $b = 2.5$, and $c = 4.5$. The above algorithm starts with the initial values, $\boldsymbol{\theta}_0$ and σ_0^2, set equal to their least-squares (LS) solutions, i.e.,

$$\boldsymbol{\theta}_0 = \boldsymbol{\theta}_{\mathrm{LS}} = \left[\sum_{s\in\Omega} \mathbf{z}_s \mathbf{z}_s^{\mathrm{T}}\right]^{-1} \left[\sum_{s\in\Omega} \mathbf{z}_s z_s\right] \tag{13a}$$

$$\sigma_0^2 = \sigma_{\mathrm{LS}}^2 = \sum_{s\in\Omega} (z_s - \boldsymbol{\theta}_0^{\mathrm{T}} \mathbf{z}_s)^2 / N_\Omega, \tag{13b}$$

where N_Ω is the number of pixels in the image region, i.e., it is M^2 for the square window indicated early on in this section.

It has been shown [19] that the above robust algorithm will converge if it starts with proper initial values. A possible candidate is the LS estimate. The iteration is stopped when the maximum of the absolute differences in the parameter estimates found by two consecutive iterations of the algorithm is equal to or less than 0.0001. In the experimental section we found the typical number of iterations to be 3 to 5, with a maximum of 8 encountered.

V. CORRELATION-CLASSIFICATION DEFECT RECOGNITION

In applying the proposed texture analysis method to wood defect recognition, an 8-pixel symmetrical neighborhood system or prediction mask for the center pixel (0, 0) is chosen as {(+1, 0), (0, +1), (−1, 0), (0, −1), (−1, −1), (−1, +1), (+1, −1), (+1, +1)} [15]. Eleven feature values, including the eight elements of $\boldsymbol{\theta}$, an estimate of the error variance σ_{k+1}^2 at convergence, the image sample mean μ_y, and the image sample variance σ_y^2 are chosen to construct a feature vector for each candidate defect region. Having chosen features, we can adopt a Bayesian or other standard rule for classification. A minimum distance classifier [17] is employed because of its simplicity and low computational requirements, which is especially important in the later stages of the vision system. In the proposed scheme, the distance between a *candidate feature vector* and a *standard feature vector* is defined as the difference of 1.0 and the maximum of the circularly shifted correlation coefficient vector [21]. The more a candidate feature vector is correlated with a standard feature vector, the shorter is the distance.

For texture pattern training representative rectangular windows in a *standard orientation* are selected from each class of defect images, and their feature values are estimated and averaged to create the standard feature vector against which a candidate feature vector associated with an image region is to be compared. Since there exists a strong directionality in the wood-grain process and the selected rectangular windows may have varying orientations with respect to a standard orientation, each candidate feature vector is circularly shifted until the maximum correlation is achieved between it and the standard feature vectors. Texture discrimination is then performed by examining the maximum correlation coefficient, and the candidate texture is classified as the texture type for which it has the highest correlation.

The defect recognition scheme can be described by two steps, *partial correlation* and *full correlation*. Since the only components of a feature vector that are direction-dependent are the eight SAR model parameters, a partial correlation with circular shifting is performed first between the first eight feature values of a candidate vector and the standard vectors. Second, a full correlation is calculated to include the three remaining features, i.e., the estimated error variance, the image sample mean, and the image sample variance. For the 11-bit image data used in this study, the first quantity of these three remaining features has a fairly small value, whereas the other two quantities have relatively large values. Therefore, these three feature values need to be properly scaled by two different weighting constants α_1 and α_2 so that they are given approximately the same weight as the eight model parameters. In particular, the value of the estimated error variance is scaled by α_1, while those of the image sample mean and variance are scaled by α_2.

Let us denote the properly weighted standard feature vector and candidate feature vector by $\mathbf{u}$ and $\mathbf{v}$, respectively, with $\mathbf{u} = [u_1, u_2, \ldots, u_N]^{\mathrm{T}}$, $\mathbf{v} = [v_1, v_2, \ldots, v_N]^{\mathrm{T}}$, where N is the total number of features selected for recognition. The elements of $\mathbf{u}$ are derived from the standard features as follows:

$$u_i = \theta_i; \quad i = 1, 2, \ldots, p \tag{14a}$$

$$u_{p+1} = \alpha_1 \sigma^2 \tag{14b}$$

$$u_{p+2} = \alpha_2 \mu_y \tag{14c}$$

$$u_{p+3} = \alpha_2 \sigma_y^2. \tag{14d}$$

The elements of $\mathbf{v}$ are defined similarly, based on candidate features.

Now the full correlation, denoted by $FC(k; \mathbf{u}, \mathbf{v})$, is defined as

$$FC(k; \mathbf{u}, \mathbf{v}) = \frac{\sum_{i=1}^{p} u_i v_{1+[i+k-1]} + \sum_{i=p+1}^{N} u_i v_i}{\sqrt{\sum_{i=1}^{N} u_i^2} \sqrt{\sum_{i=1}^{N} v_i^2}}; \quad k = 0, 1, \ldots, p-1, \quad (15)$$

where $N = p + (N - p)$, with p the number of estimated SAR model parameter features and $(N - p)$ the number of weighted features. The index $[x]$ denotes the modulo-p value of its argument, as necessary to evaluate a circular correlation. Having defined $FC(k; \mathbf{u}, \mathbf{v})$ as in equation (15), the *distance* $d(\mathbf{u}, \mathbf{v})$, between vectors $\mathbf{u}$ and $\mathbf{v}$, can be expressed as

$$d(\mathbf{u}, \mathbf{v}) = 1 - \max_{k}\{FC(k; \mathbf{u}, \mathbf{v})\}. \quad (16)$$

This value $d(\mathbf{u}, \mathbf{v})$ is evaluated for the set of possible standard feature vectors $\mathbf{u}_l$ (i.e., for each defect; $l = 1$ for Bark, $l = 2$ for Decay, and $l = 3$ for Knot), and then forwarded to the minimum distance correlation-classifier

$$l^* = \arg\min_{l}\{d(\mathbf{u}_l, \mathbf{v})\} \quad (17a)$$

which classifies the weighted candidate feature vector $\mathbf{v}$ as being associated with the defect l^*, one of the prototypical defects. Note that

$$l^* = \arg\max_{l}\{\max_{k}\{FC(k; \mathbf{u}_l, \mathbf{v})\}\} \quad (17b)$$

is an alternative expression for the minimum distance correlation-classifier.

VI. EXPERIMENTAL RESULTS

Hardwood logs were selected from a local log yard in Virginia. They were transported to a local hospital in Blacksburg, VA, where a Siemens medical CT-scanner was used to generate log CT images containing 256 × 256 pixels. The gray levels of these 11-bit log images were from – 1024 to 1024. With an interstice resolution of 0.8 mm, every pixel of the images represented an area of 0.25 × 0.25 mm of the cross-section of the logs. Two sets of experiments on defect recognition were conducted with some of these image samples available from the Wood Image Library at the Spatial Data Analysis Laboratory. All processing was run on a VAX11/785 system. Log CT images were first segmented by the low level module of the vision system [5], and then the segmented images were forwarded to the high level module where defect recognition was performed using the above-mentioned approach.

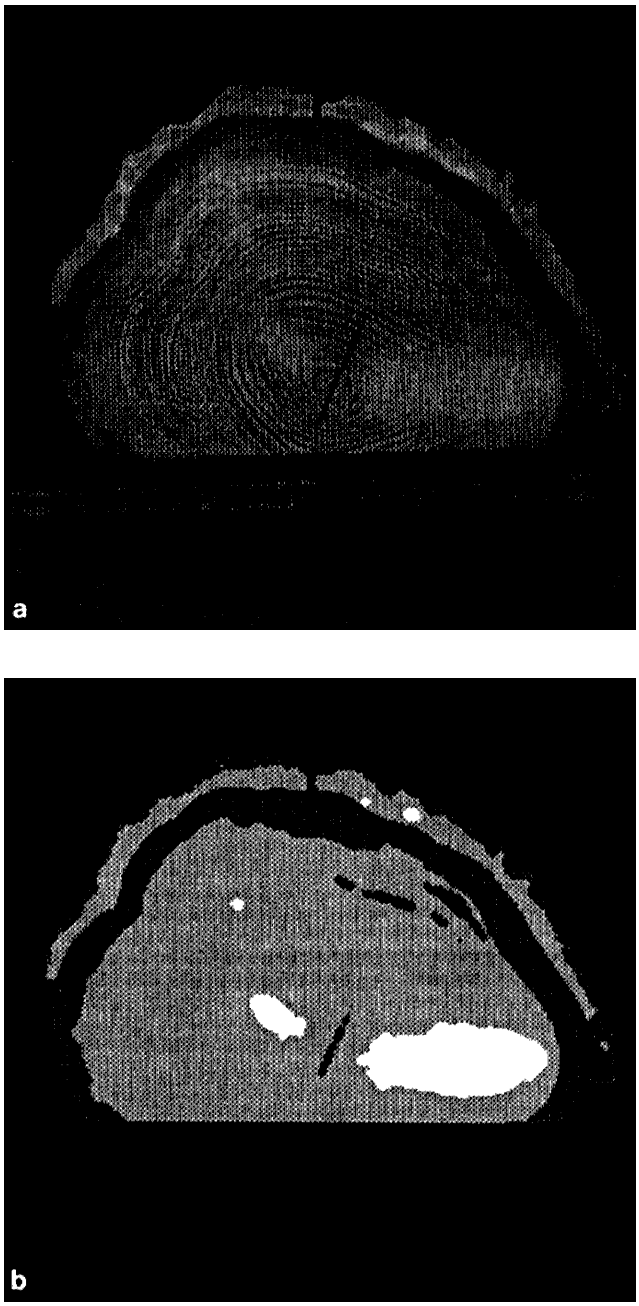


FIG. l(a). Image RK 11.22 with Bark and two Knots. (b) The segmentation of image RK 11.22.

The experimental results for the proposed method were obtained with CT images of two different red oak trees. Figures 1 and 2 each show one slice of these two logs. The images are segmented and different gray-level regions are assigned arbitrary labels to signify different wood defects in the log. To make the segmentation graphically discernible, a pseudo-color is assigned to

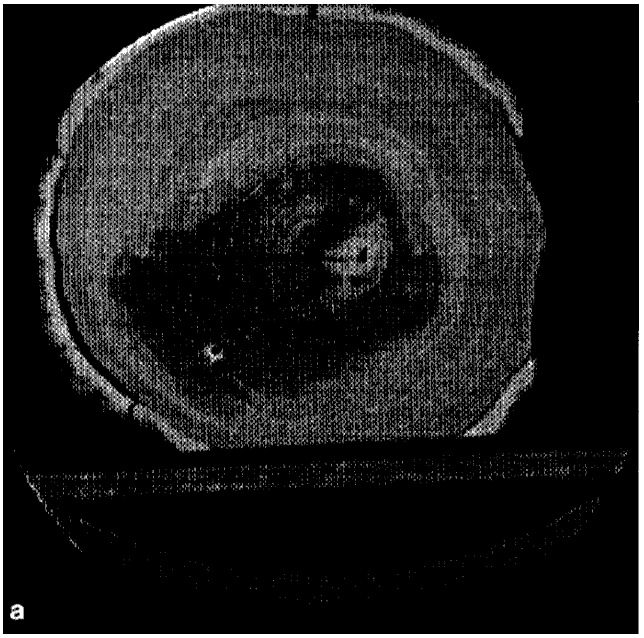


FIG. 2(a). Image RK 12.05 with Decay and Stains. (b) The segmentation of image RK 12.05.

each segment. Due to different red oak wood samples, the same type of wood defect may be marked with different pseudocolors on the two segmented images. Figure la is the 22nd slice from a sequence of images for the first red oak log. Figure 1b exhibits its segmentation which contains barks, knots, a split, and clear-wood; the latter occupying most of the image area. Figures 2a and 2b are the corresponding images for the 5th slice from the second red oak tree, containing a brown decay as well as barks. Both Figs. 1b and 2b represent correct segmentation of the bark, while different pseudocolors were assigned to the bark on those two images.

Using the model and algorithm described previously, the features of each defect were estimated from several image slices and their values averaged, to produce a standard feature vector for each defect type. A CT image was 256 × 256 × 11 bits with gray-levels ranging from 0 to 2047 by a gray-level transformation. After image segmentation [5], windows of appropriate sizes and orientations were selected for each potential defect region. The original image data within such windows were standardized by removing the mean value and then scaling by the standard deviation. First, causal 3-pixel and symmetric 4-pixel neighborhood systems were tried as the SAR model masks. Since each wood-grain is highly correlated with all its neighbors, these *partial* model masks did not produce consistent parameter estimates for the textures under study. Next the 8-pixel neighborhood system ($p = 8$) was selected as the SAR model mask. Eleven feature parameters were estimated, using the proposed iterative estimation algorithm. These parameters were then cross-correlated with each of the standard vectors, and the calculated correlation values were input to the minimum distance classifier to determine the defect type of a candidate region.

In the *first experiment,* 7 bark regions from 2 image slices (slices 20 and 24 of red oak section 11), 7 decay regions from 1 slice (slice 1 of red oak section 12), and 7 knot regions from 2 slices (slices 21 and 22 of rcd oak section 11) were selected as a training set and used to create the standard feature vector. Correlation-classification was then conducted with the training set, and the resulting classification accuracies were 71.4%, 100%, and 100% for decays, barks, and knots, respectively. A testing set of 15 bark, 21 decay and 13 knot candidate regions was selected next, from 14 different image slices (including those used in the training), and correlation-classification conducted by the above described method. To achieve stable solutions, the step size of the iterative algorithm was selected as = 0.01. As the typical value of the estimated error variance is on the order of 10^{-1}, whereas those of the image sample mean and variance are from 10^{3} to 10^{5}, the weighting constants were chosen as $_1$ = 100, and $_2$ = 0.0001 so that all features are treated approximately equally by the classifier. The following defect recognition accuracies were obtained with the testing set: 80% for bark regions with 3 out of 15 barks misclassified as decays; 81% for decay regions with 4 out of 14 decays misclassified as knots; and 100% for knot regions with no misclassification.

Since the above test set includes the training samples used in creating the standard feature vectors, this raises the apparent success rate of the experiment. To demonstrate the practical success rate better, an experiment

TABLE 1
Correlation-Classification Result on Pure Test Samples

Texture sample (defect region)	$\max_k FC(k; \mathbf{u}_l, \mathbf{v})$			Classification type (X ~ wrong)
	l - Bark	l - Decay	l - Knot	
Bark 1	0.9940	0.7103	0.6125	Bark
Bark 2	0.9701	0.5236	0.4002	Bark
Bark 3	0.9275	0.4017	0.2592	Bark
Bark 4	0.9667	0.5612	0.4581	Bark
Bark 5	0.9812	0.5677	0.4432	Bark
Bark 6	0.8000	0.9803	0.9453	Decay (X)
Bark 7	0.9512	0.8976	0.8088	Bark
Decay 1	0.9379	0.9230	0.8511	Decay
Decay 2	0.8643	0.9609	0.9120	Decay
Decay 3	0.8223	0.9801	0.9413	Decay
Decay 4	0.7295	0.9991	0.9810	Decay
Decay 5	0.6908	0.9990	0.9910	Decay
Decay 6	0.6606	0.9990	0.9940	Decay
Decay 7	0.6532	0.9980	0.9956	Decay
Decay 8	0.6231	0.9950	0.9980	Knot (X)
Decay 9	0.6742	0.9991	0.9944	Decay
Decay 10	0.6037	0.9920	0.9990	Knot (X)
Decay 11	0.6533	0.9970	0.9960	Decay
Knot 1	0.5676	0.9853	0.9995	Knot
Knot 2	0.5692	0.9860	0.9992	Knot
Knot 3	0.5800	0.9866	0.9997	Knot
Knot 4	0.5702	0.9854	0.9985	Knot
Knot 5	0.5801	0.9878	0.9992	Knot
Knot 6	0.5682	0.9865	0.9993	Knot
Knot 7	0.5800	0.9876	0.9988	Knot
Knot 8	0.5667	0.9860	0.9986	Knot

with a pure testing data set—one which contains none of the training samples—is conducted. For this pure test, 7 bark, 11 decay, and 8 knot candidate regions were selected from several image slices (exclusive of those used in the training step). Table 1 provides the correlation-classification results for the pure test. Summary of the results shows that the correlation-classification algorithm produced the following correct defect recognition percentages: 86% for bark regions, 82% for decay regions, and 100% for knot regions.

We note that the correlation-classifier, as used, produces too many erroneous results. Certainly one expects the classification performance for the learning set to be excellent, with the hope of having designed a classifier that is not very sensitive to the change of using candidate features rather than standard features. The latter, practical situation, is expected to cause some deterioration in performance, so that the aim is to mitigate that deterioration as much as possible. From the results in Table 1, we note that the correlations for bark are generally much higher, or much lower, than for either decay or knot. The latter 3 types of wood defects, as studied here, can be subdivided into 2 main groups: the outer wood group (consisting of bark), and the inner wood group (consisting of decay and knot). The correlation-classification method did not discriminate correctly between the inner and outer wood groups for Bark-sample 6. For Decay-sample 8 and Decay-sample 10 the correlations are both very high and different only in the third decimal place. The latter suggests that a single feature, which is similar for both Decay- and Knot-samples, is dominating the correlation. This is possibly exacerbated by the fact that the two scaling constants $_1$ and $_2$ were not optimized for maximum separability between the three classes of defects.

In the *second experiment*, a two-level tree classifier structure was employed. The first level classifies candidates into the inner *(I)* and the outer *(O)* defect groups by examining the value of the estimated error variance *(EEV)* according to the following threshold test:

$$\sigma^2 \underset{O}{\overset{I}{\gtrless}} \tau. \tag{18}$$

The *EEV* threshold τ used was 0.08. The selection of the single *EEV* texture feature for the first level classification is equivalent to a feature space transform, wherein an optimal subset of features is chosen from the feature vector in order to increase the separability of the inner and outer defect classes [17]. The second level classifies according to the correlation-classifier using the remaining (here, ten) features. The latter classification has been reduced to the classical pair-wise Bayesian problem, in which the distance measure between two vectors is used to discriminate decay from knots. The two-level correlation-classifier is depicted in Fig. 3. The results of using this two-level correlation-classifier, on the same pure testing data as used for the first experiment, are given in Table 2. Note that the level-1 estimated error variance feature easily discriminates between the inner and outer

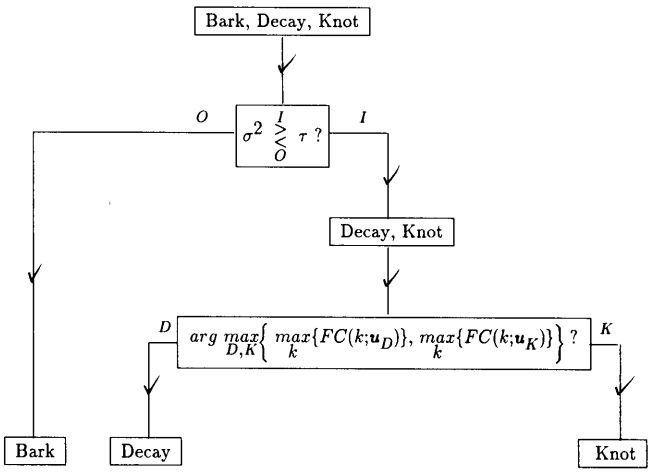


FIG. 3. The two-level correlation-classifier.

defect groups, including for Bark-sample 6. Also, the level-2 correlation-classifier now correctly classifies Decay-samples 8 and 10. For the two-level correlation-classifier we thus report a 100% classification accuracy for each of the three groups of defects; in fact, this holds for both the training and the testing data sets described under the first experiment.

Lest one believes that the two-level classifier based on the correlation-classifier results in a perfect classifier, it should be noted that, however many experiments one has performed in the past, a future sample may come along that defies classification. Instead, we conservatively state that with the two-level correlation-classifier, which was arrived at via the "optimization" of choosing a single feature for its first level, the overall results have improved considerably.

Given the CT image data of the wood species available, the current vision system for hardwood log inspection can handle such prototypical defects as barks, decays, and knots. As CT image data for more species becomes available the vision system could be enhanced to classify other defects as well, such as stains, mineral streaks, holes, etc.

VII. CONCLUDING REMARKS

In this paper, a stochastic field-based approach for wood texture analysis was presented. In particular, an iterative robust image modeling algorithm was applied to the problem of defect recognition in a computer vision system for log inspection. Experiments show that the described texture modeling method appears capable of discriminating between clear-wood and several defects, including knots, decays, and bark. Experiments also indicate that the causal 3-pixel neighborhood and the symmetrical 4-pixel neighborhood configurations are not appropriate for modeling wood-grain textures; that textures of the same grain, due to the wood growth process, demonstrate scale- and orientation-variation at different locations on each CT log image; and that many image preprocessing methods (histogram equalization, adaptive noise smoothing, or pixel interpolation in the circular model [15]) tend to adversely affect the discrimination power of the described texture modeling method.

Although there have been several research papers on texture invariance problems, a more practical method is needed for a robust approach to modeling real-world wood textures. In future research, the initial value problem of the iterative algorithm needs to be studied further, and solutions like the L_1 and the high-breakdown estimates could be selected as the initial values [22]. Moreover, more robust methods need to be developed to compute the covariance matrix (related to the weight $W1(s)$). Yet another avenue of research is to incorporate features other than texture-related ones, such as a region's geometric properties (orientation, compactness, elongatedness). This is an effort toward an efficient object recognition scheme as an integral part of a vision system that can automatically identify different defects in a variety of log images.

TABLE 2
Two-Level Classification Result on Pure Test Samples

Texture sample (defect region)	EEV	$\max_k\{FC(k; \mathbf{u}_l, \mathbf{v})\}$		Classification type (X ~ wrong)
		l - Decay	l - Knot	
Bark 1	0.0296			Bark
Bark 2	0.0517			Bark
Bark 3	0.0092			Bark
Bark 4	0.0183			Bark
Bark 5	0.0233			Bark
Bark 6	0.0612			Bark
Bark 7	0.0325			Bark
Decay 1	0.3415	0.9302	0.8651	Decay
Decay 2	0.1279	0.9569	0.9201	Decay
Decay 3	0.1543	0.9766	0.9425	Decay
Decay 4	0.3886	0.9865	0.9353	Decay
Decay 5	0.2107	0.9976	0.9886	Decay
Decay 6	0.1443	0.9989	0.9943	Decay
Decay 7	0.2114	0.9974	0.9932	Decay
Decay 8	0.2557	0.9920	0.9806	Decay
Decay 9	0.1023	0.9993	0.9967	Decay
Decay 10	0.3011	0.9995	0.9913	Decay
Decay 11	0.2356	0.9993	0.9976	Decay
Knot 1	0.1134	0.9765	0.9993	Knot
Knot 2	0.3353	0.9889	0.9994	Knot
Knot 3	0.5178	0.9794	0.9969	Knot
Knot 4	0.1667	0.9903	0.9992	Knot
Knot 5	0.1923	0.9857	0.9990	Knot
Knot 6	0.1748	0.9877	0.9991	Knot
Knot 7	0.1652	0.9799	0.9989	Knot
Knot 8	0.1678	0.9937	0.9960	Knot

ACKNOWLEDGMENTS

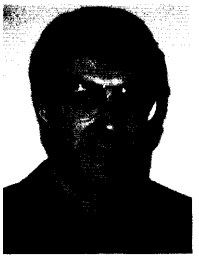
We express our appreciation to Mr. Phillip A. Araman of the USDA Forest Service Southeastern Forest Products Experiment Station for his invaluable support for the hardwood log inspection project. The first author also thankfully acknowledges the help with the iterative robust estimation algorithm provided by Dr. L. Mili of the Bradley Department of Electrical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

REFERENCES

1. T-H. Cho, R. Conners, and P. Araman, A computer vision system for automatic lumber detection using blackboard expert system method, in *Proc. 10th International Conference on Pattern Recognition, Atlantic City, NJ, June, 1990.*
2. R. W. Conners, C. W. McMillin, K. Lin, and R. E. Vasquez-Espinosa, Identifying and locating surface defects in wood: Part of an automated lumber processing system, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-5** (6), Nov. 1983, 573-583.

3. B. V. Funt and E. C. Bryant, A computer vision system that analyzes CT-scans of sawlogs, in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition, 1985,* pp. 175-177.
4. A. Koivo and C. Kim, Robust image modeling for classification of surface defects on wood boards, *IEEE Trans. Systems Man Cybernet.* **SMC-19** (6), Nov/Dec 1989, 1659-1666.
5. D. Zhu, R. Conners, and P. Araman, Image Processing for hardwood log inspection, *SPIE Conf. on Applications of Digital Image Processing XIV, San Diego, CA, 1991.*
6. F. Wagner, F. Taylor, D. Ladd, C. McMillin, and F. Roder, Ultrafast CT scanning of an oak log for internal defects, *Forest Products J. 39,* Nov/Dec 1989.
7. D. G. Hodges, W. C. Anderson, and C. W. McMillin, The economic potential of CT scanners for hardwood sawmills, *Forest Products J.* **40** (3), Mar. 1990, 65-69.
8. F. Wagner, F. Taylor, P. Steele, and T. E. Harless, Benefit of internal log scanning, in *Proc. 3rd international Conference on Scanning Technologies in Sawmilling, October 1989,* pp. 1-16.
9. L. G. Occena and J. M. A. Tanchoco, Pattern directed extraction and characterization of defect configuration in solid log models, *Artif. Intell. Engrg,* **4** (3), 1989, 144-154.
10. R. Chellapa and S. Chatterjee, Classification of textures using Gaussian Markov random fields, *IEEE Trans. ASSP* **ASSP-33** (4), Aug. 1985, 959-963.
11. Z. Lin and Y. Attikiouzel, 2-d linear prediction model-based decorrelation method, *IEEE Trans. Pattern Anal. Mach. Intell.* **11** (6), June 1990, 661-665.
12. R. M. Haralick, Statistical and structural approaches to textures, *Proc. IEEE* **67** (5), 1979, 786-804.
13. R. Conners and C. Harlow, A Theoretical comparison of texture algorithms, *IEEE Trans. Pattern Anal. Mach. Intell.* **2** (3), May *1981, 204-222.*
14. R. Bajcsy and L. Lieberman, Computer description of real outdoor scenes, in *Proc. 2nd Int'1 Joint Conf. on Pattern Recognition, August 1974,* pp. 174-179.
15. R. L. Kashyap, Image models, in *Handbook of Pattern Recognition and Image Processing,* Academic Press, New York, 1986.
16. J. Zhang and J. W. Modestino, A model-fitting approach to cluster validation with application to stochastic model-based image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAM1-12** (10), Oct. 1990, 1009-1017.
17. R. Duda and P. Hart, *Pattern Classification and Scene Analysis,* Wiley, New York, 1973.
18. J. C. Bezdek, *Pattern Recognition with Fuzzy Objection Function Algorithms,* Plenum Press, New York, 1981.
19. P. J. Huber, *Robust Statistics,* Wiley, New York, 1981.
20. L. Ljung and T. Söderström, *Theory and Practice of Recursive identification,* The MIT Press, Cambridge, MA, 1983.
21. D. Zhu, R. Conners, and A. A. Beex, Parameter modeling of CT image textures for wood defect recognition, in *Proceedings of 25th Conference on Information Sciences and Systems, Johns Hopkins University, March 1991.*
22. P. J. Rousseeuw and J. Leray, *Robust Regression and Outlier Detection,* Wiley, New York, 1987.

DONGPING ZHU obtained a B.S. in radio electronics and an M.S. in communications systems in 1982 and 1985, respectively, both from Dalian University of Technology (formerly Dalian Institute of Technology), Dalian, P.R. China. From 1985 to 1987, he was with the Department of Electrical Engineering at DUT as a research associate and instructor. In 1989, he earned an M.S. in electrical engineering from Virginia Polytechnic Institute and State University, where he has since been working toward a Ph.D. in electrical engineering at the Spatial Data Analysis Laboratory. Since January 1992, he has been employed as Staff Scientist with InVision Technologies, Inc., Foster City, California. He is a member of IEEE, SPIE, and IAPR. His major research interests include stochastic signal and image processing, system theory, computer vision, pattern recognition, uncertainty reasoning in AI, and their industrial applications.



A. A. (LOUIS) BEEX received the "Ingenieur" degree in 1974, from Technical University Eindhoven, the Netherlands. His Ph.D. degree, in 1979, was from Colorado State University, Fort Collins, CO. Both are in electrical engineering. From 1976 to 1978 he was a staff research engineer at Starkey Laboratories, Minneapolis, MN, working on digital signal processing applications. Since 1979 he has been on the faculty of the Bradley Department of Electrical Engineering at Virginia Tech, Blacksburg, VA. He has served as a consultant to Colorado State University, Fort Collins, CO, to the Naval Surface Weapons Center, Dahlgren, VA, and to Rome Air Development Center, Rome, NY. He is past associate editor of the *IEEE Transactions on Acoustics, Speech, and Signal Processing.* His research interests lie in stochastic and digital signal processing, such as, more specifically, robust spectral analysis and modeling, and its applications, image representation, modeling, coding, and reconstruction, sensor array processing, sensitivity analysis of digital filters, and system identifiers.