Nondestr. Test. Eval., Vol. 15, pp. 279-309

NONDESTRUCTIVE EVALUATION OF HARDWOOD LOGS: CT SCANNING, MACHINE VISION AND DATA UTILIZATION

DANIEL L. SCHMOLDT[a,*], LUIS G. OCCEÑA[b],
A. LYNN ABBOTT[c] and NAND K. GUPTA[d]

[a] *USDA Forest Service, Biological Systems Engineering Dept., 460 Henry Mull, University of Wisconsin, Madison, WI 53706-1561, USA;* [b] *Department of Industrial and Manufacturing Systems Engineering, University of Missouri, Columbia, MO 65211, USA;* [c] *Bradley Department of Electrical Engineering, Virginia Tech, Blacksburg, VA 24061-0111, USA;* [d] *Omega Technologies International, Lakemoor, IL 60050-8653, USA*

(Received 20 August 1998)

Sawing of hardwood logs still relies on relatively simple technologies that, in spite of their lack of sophistication, have been successful for many years due to wood's traditional low cost and ready availability. These characteristics of the hardwood resource have changed dramatically over the past 20 years, however, forcing wood processors to become more efficient in their operations. In spite of some recent advances, the breakdown of hardwood logs into lumber continues to be hampered by the inability of sawyers to "see" inside of the log prior to making irreversible cutting decisions. The need for noninvasive assessment of hardwood logs prior to breakdown is well accepted, but is difficult to realize because industrial scanning. in this context, is unique in several respects. For example, large volumes of material must be inspected quickly over an extended duty cycle, the wood material still possesses relatively low value compared to other industrial materials that require internal scanning, and many wood processors are small operations located in rural areas. Successful implementation of new scanning technology, however, will have tremendous payback for wood processors. and for timber resource conservation efforts. The research program reviewed here applies a three-pronged approach to address this situation. First, a relatively new and innovative CT scanning technology is being developed that can scan hardwood logs at industrial speeds. Second. machine vision software has been created that can interpret scanned images rapidly and with high accuracy. Third, we have developed 3-D rendering and analysis techniques that will enable sawyers to apply image assessment to actual log

*Corresponding author. E-mall: dlschmol@facstaff.wisc.edu.

breakdown. This integrative research direction combines hardware and software systems to scan logs, process images, and apply imaging to real-time, industrial decision-making.

Keywords: Computed tomography; Wood utilization; Log scanning; Automated processing

INTRODUCTION

The manufacture of furniture, cabinets, flooring, millwork, and molding, along with hardwood exports, accounts for most of the high- and medium-grade hardwood lumber consumption in the US [1]. In contrast to softwood lumber, which is valued in terms of volume and mechanical strength, the value of hardwood lumber is based more heavily on appearance-related criteria. The conversion of hardwood trees into final commercial products involves a number of steps (Fig. 1). First, tree-length material is "bucked" into logs in the forest; these logs are subsequently converted to lumber in sawmills. For the most

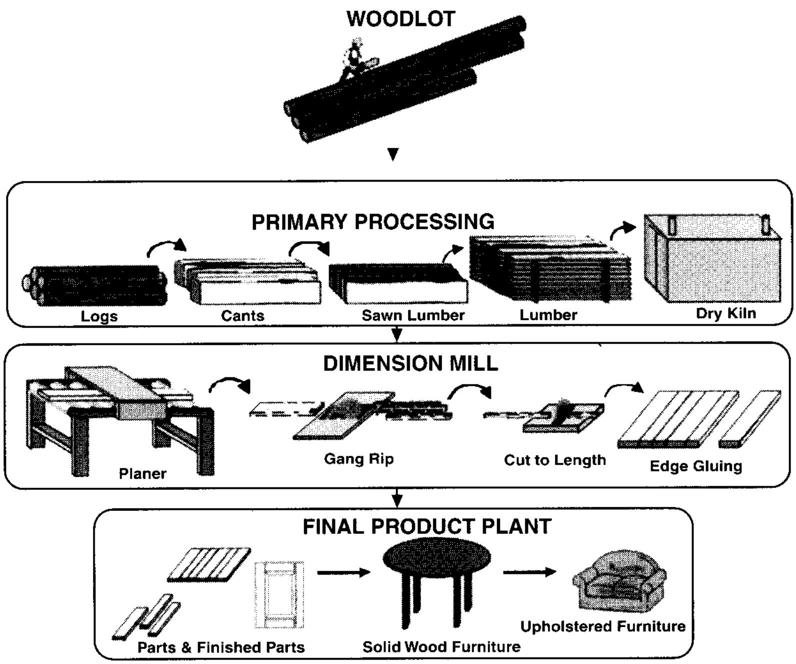

FIGURE 1 The hardwood processing industry consists of 4 segmented processing stages: log bucking, sawmills, dimension mills and manufacturing plants.

part, only "clear wood" (defect-free) portions of each board can be used in finished products; therefore, dimension mills cut, chop, and mould the wood into small usable parts of pre-determined dimensional sizes. In the final step, these parts are glued, assembled, and painted/stained to produce the desired finished wood products. Although processing integration is reviewed in the last section, the remainder of this paper focuses primarily on processing of logs into lumber (often called primary processing).

In a typical hardwood sawmill, logs enter the mill and go through a de-barking process (Fig. 2). Following this operation, they go to the headrig where a sawyer moves the log repeatedly past a saw to remove boards one at a time. As more of a log's interior is exposed with the removal of each board, the sawyer may re-orient the log periodically to cut from the best side, or to restrict a log's defects to the minimum number of boards (or the edges of those boards). Sawn boards go through subsequent operations of edging and trimming, where defects near the edges and/or ends of the boards are removed to increase each board's grade, and therefore its commercial value. The cant (the center section of the log, which appears rectangular in cross-section), remaining from initial breakdown, may either (1) enter a resawing operation where additional boards are cut, or (2) be sold intact for use in low-value products, such as pallet material. Cant material is of, generally, low value because (1) many of a tree's branches begin there, giving rise to knots in the wood, (2) it contains the tree's pith, which is composed entirely of soft tissue, (3) fungal infection from the roots

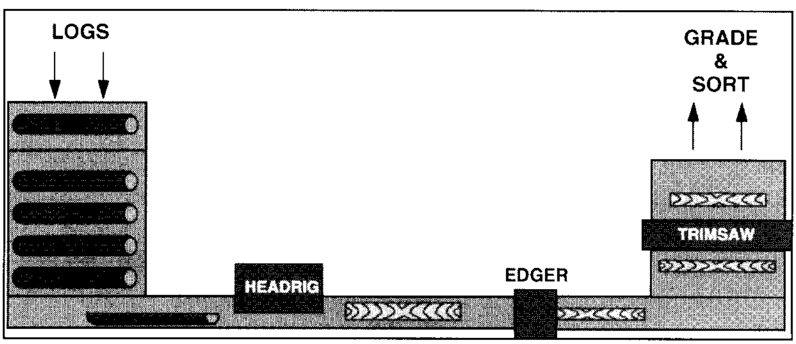

FIGURE 2 A typical sawmill contains debarking, log sawing. edging/trimming and grading/sorting operations.

often results in decay spreading vertically up through the tree's core, and (4) the enter of a log has greater curvature in each annual ring, which produces much greater drying stresses in the resulting boards.

Because the value of hardwood lumber is heavily dependent on the quantity, type, and location of defects, each log must be sawn to minimize (subject to board sizing constraints) the number, size, and severity of defects in the resulting boards. During hardwood log breakdown, profit-critical decisions are made by the sawyer that can significantly affect downstream processing operations. This observation suggests that targeting sawlog breakdown improvements can drastically increase lumber value recovery. Traditionally, the sawyer chooses a sawing strategy by visually examining the exterior of the log, modifying the strategy as sawing exposes the log interior. The sawyer uses log shape, external indicators of internal defects, and knowledge of lumber grades to make sawing decisions. While sawyers are highly skilled in this task, studies [2–4] have shown that the lumber value of logs can be improved 20% or more by carefully selecting the proper sawing strategy. However, the current level of information available to sawyers during the log breakdown operation is inadequate for enhancing the sawyer's capability to produce high-value boards. Developing nondestructive sensing and analysis methods that can accurately detect and characterize interior defects is critical to future efficiency improvements for sawmills [5].

Because most defects of interest are internal, a nondestructive sensing technique is needed that can provide a 3-D view of a log's interior. Several different sensing methods have been tried, including nuclear magnetic resonance [6], ultrasound [7] and X-rays [8–13]. Due to its efficiency, resolution, and widespread application in medicine, X-ray computed tomography (CT) has received extensive testing for roundwood applications [11,13–17]. An X-ray CT scanner produces image "slices" that capture many details of a log's internal structure (Fig. 3). Because X-ray attenuation is linearly related to wood density [18] and many wood features (including defects) exhibit density differences [19], many lumber-quality defects (e.g., knots, voids and decay) are apparent in CT images.

While economic analyses suggest that lumber value gains can offset scanning costs [20,21], there are several technological hurdles that must be overcome for the application of computer tomography scanning to

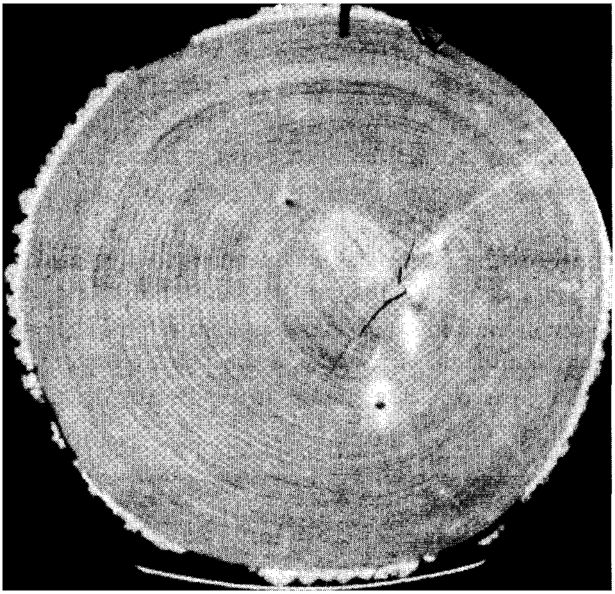

FIGURE 3 This 512 × 512 black cherry CT image depicts the density variation possible in tomographs.

sawlogs. First, one must determine what characteristics are required from an industrial scanner in order to adequately image logs with variation in size and species. In addition, spatial resolution requirements and levels of image contrast will vary between logs. Second, there must be a way to condense the tremendous amounts of data that are generated by CT imaging, so that only information critical for decision-making is retained for downstream processing. Finally, the CT data need to be visualized in a way that conveys their spatial nature and that is natural for the sawyer to understand [22]. These issues constitute our three-pronged research program; each is reviewed in the following sections.

CT IMAGING OF HARDWOOD LOGS

Several capabilities are essential for application of CT imaging in hardwood sawmills [22]. These include the ability to scan large diameter logs, to provide relatively high-resolution images, to perform scans quickly, and to scan logs for long production shifts. Although medical CT systems would appear to be easy to adapt to log scanning, it is, unfortunately, the case that current medical CT systems have been engineered for low frequency, short duration use. This is incompatible with industrial sawmill needs. Direct application of existing medical scanning technology would be, in most cases, prohibitively expensive and slow. Rather, an industrial CT scanner needs to be used. Nevertheless, most existing industrial CT scanners are designed for quality control inspections in off-line situations, or on-line where materials are relatively small and of limited mass (e.g., airline baggage inspection). Industrial scanning of large-volume and -mass objects (e.g., logs) in an on-line operation demands that we investigate alternative CT technologies.

Existing CT Technology

Current CT scanner technology includes four types of scanner systems, referred to as “generations”. They are of two basic types: (1) parallel and (2) fan X-ray beam scanners. There are two types of parallel X-ray beam scanners: first- and second-generation systems. Also, there are two types of fan X-ray beam scanners: third- and fourth-generation systems. The following subsections briefly describe each one and identify known strengths and weaknesses.

First-generation Scanners

First-generation CT scanners use a single X-ray detector (Fig. 4(a)). A pencil X-ray beam is formed by the X-ray source and the detector. This X-ray beam is traversed over the scanned object to measure the X-ray intensities through parallel paths in the object. A complete set of such measurements is made through the entire extent of the object (from one edge to the other edge). After each such complete set of measurements, the object is rotated by a small angle (typically by 1° between views) and the parallel measurement process is repeated. Scanning is continued until measurements have been made through 180° of view angles.

First-generation systems possess a number of strengths owing to their design simplicity. These include: (1) low expense, (2) simple data

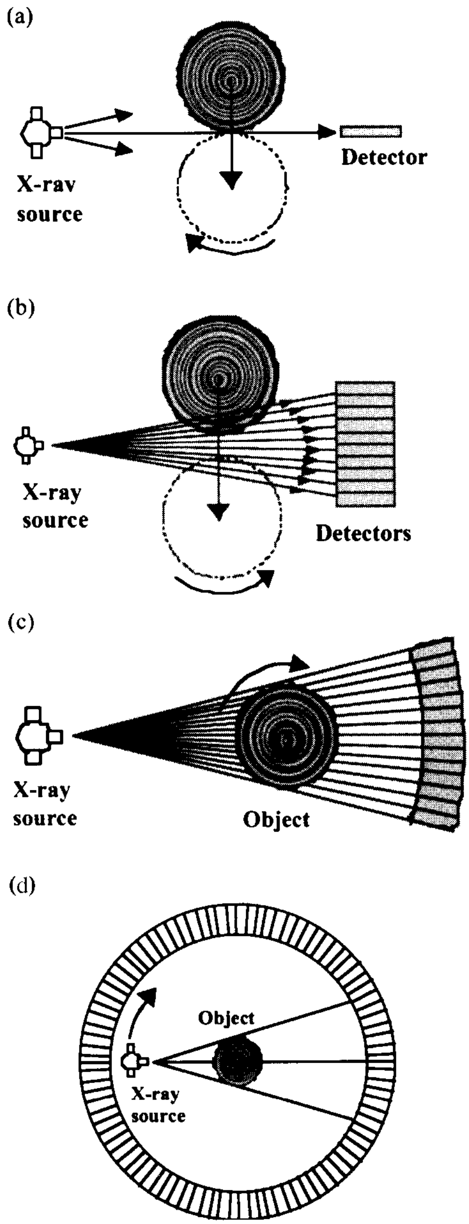

FIGURE 4 First (a), second (b), third (c) and fourth-generation (d) CT geometries are shown. In each case, the detectors are perpendicular to the axis of rotation, so that scanning creates an axial tomograph.

collection scheme because of the one detector, (3) parallel X-ray beam data collection requires relatively simple algorithms to reconstruct tomographs, (4) no multiple-detector mismatches that often lead to image noise, and (5) any size object can be scanned by adjusting the traverse length of the pencil X-ray beam. Despite their advantages, however, first-generation CT scanners are prohibitively slow for most applications where time is a critical parameter, therefore, they are almost never used.

Second-generation Scanners

Second-generation CT scanners use a detector system (array) consisting of several X-ray detectors (Fig. 4(b)). The X-ray detectors form independent pencil beams – at slightly different angles – with the X-ray source. The detector system makes simultaneous measurements through different angles in a single traverse. After a set of these simultaneous measurements through the entire extent of the object, the object is rotated by the array beam's angle and the measurement process is repeated again until the sequence of rotation and transversal collects 180° data.

Second-generation systems possess most of the advantages of first-generation systems, including simple geometry and data collection scheme, easy reconstruction algorithms, and unlimited object sizes. In addition, multiple detectors can collect data simultaneously, so fewer traverses are required. Second-generation systems also suffer from excessive down time needed for mechanical operations, multiple image traverses, and single-slice data collection. Furthermore, the following disadvantages also exist. First, several detectors are used to collect the data for a single tomograph, which means that there can be, and usually are, variations between the response of various detectors. Even following software corrections, a small amount of additional noise is added to the data, resulting in a small loss of image quality. Second, small artifacts appear in reconstructed CT images due to small mismatches in the data from various detectors. Third, to collect a complete set of data through all angles in the object, the inside edge of the X-ray fan beam must touch the outer surface of the object at the beginning, as well, at the end of each traverse. Hence, a significant amount of useless data is collected at the beginning and end of each traverse.

Third-generation Scanners

Third-generation CT scanners use a detector array with many detectors. The detectors are usually located on an arc focused at the X-ray source (Fig. 4(c)). In this case, data are collected in a fanning movement, rather than parallel. A sufficient number of detectors are used so that the fan shaped X-ray beam covers the entire object. The object (or source-detector pair) is rotated to collect the entire CT data. For 180° data, the object is rotated by 180° (plus the X-ray beam fan angle).

Third-generation systems offer several advantages over parallel-beam systems. First, data are simultaneoulsy collected through the entire object for each view. Second, the mechanical motion of the gantry is very simple rotational movement. Third, motions are continuous and hence no time is wasted in mechanical starting and stopping. Fourth, scan times are quite fast due to non-stop rotational motion and many detectors collecting simultaneous data.

At the same time, however, third-generation systems have numerous drawbacks, including the limitation of single-slice data collection. First, the maximum object diameter is limited by the number of detectors. Second, scanner resolution is fixed by the number and spacing of detectors covering the object. Third, data from all detectors are always collected. Hence, a significant amount of useless data is collected when smaller size objects are scanned. Fourth, each detector views a tangent to a fixed circle within the scanned object, causing circular artifacts in images. Fifth, system cost is high because it requires a large number of detectors to ensure coverage of large objects (without translational movement – as in second-generation scanners – a sufficient number of detectors must be installed to image the largest object).

Fourth-generation Scanners

Fourth-generation CT scanners use a detector system (array) with an even larger number of detectors. The detectors are located in a circle, which surrounds the X-ray source, and the object to be scanned (Fig. 4(d)). Because the detector array forms a circle, this system requires the greatest number of detectors. The X-ray source is located between the detector circle and the object, and is rotated in a circle to collect 180° or 360° data.

Because of their similar geometry third- and fourth-generation systems share both advantages and disadvantages. Advantages include: data are collected simultaneously through the entire object for each view, motions are continuous, mechanical motion is simple (only the X-ray source is rotated), and scan times are fast. Limitations are: the X-ray fan beam limits object size, scanning small objects results in much useless data collection, and a single slice is collected at a time. In addition, fourth-generation systems possess a very high system cost due to the large number of detectors required to cover the entire detector circle. Due to high cost, fourth-generation systems are rarely used for industrial applications (except where inspection failure losses are substantial, e.g. airport baggage explosive detection) and are becoming uncommon even in the medical industry.

Tangential Scanning Technology

Scanner Design and Operation

To overcome many of these limitations with traditional CT technology, we have examined the feasibility of using *tangential* scanning for hardwood logs [23]. In tangential scanning, the detector array is placed parallel to the axis of rotation of the object and perpendicular to the cross-section [24]. A fan shaped X-ray beam is formed by the X-ray source and the detector array and extends along the axis of rotation of the object (Fig. 5).

For data collection, the object is rotated rapidly around it's own axis. Simultaneously, the object (or source-detector movement) slowly traverses through the X-ray fan beam in a direction perpendicular to the fan beam. At the beginning of data collection, the outside surface of the object touches the X-ray fan beam. For a data set covering 180° of views, the object is traversed from its one edge to its center. For a 360° data set, the object is traversed from one edge to the other edge by the X-ray beam (or equivalently, the specimen translates).

As the object translates through the X-ray beam. the detectors collect X-ray intensity data along tangential paths of varying diameter circles. For most of the X-ray beam, each detector collects data for one cross-sectional slice of the object. In addition, only one detector collects the entire data for one cross-sectional CT slice. As one moves toward the

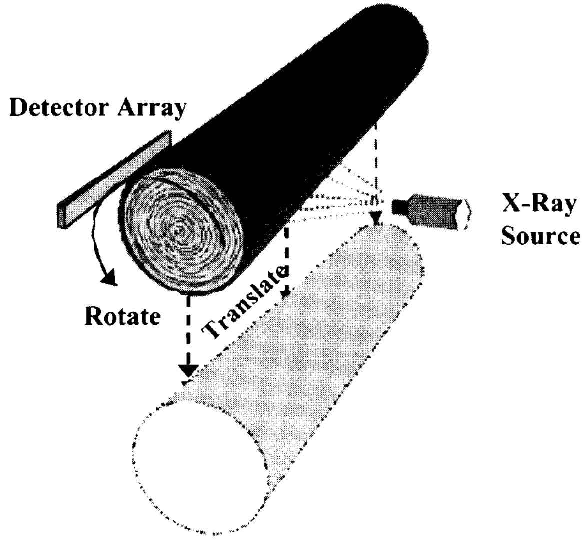

FIGURE 5 Tangential scanning geometry differs drastically from traditional scanners, wherein the specimen rotates and translates simultaneously and the detector array is parallel to the axis of rotation.

edges of the fan beam (along the s-axis of the log), however, multiple detectors collect data for a slice. For these edge slices, a 3-D-reconstruction algorithm will be needed to generate reliable tomographs. All detectors of the detector array simultaneously collect data to scan an entire sub-volume of the object. The number and spacing of detectors determine how many tomographs (and their pitch) can be collected simultaneously.

Tangential Scanning Strengths and Weaknesses

For industrial applications, the improved geometry of tangential scanning provides some important advantages over existing scanning geometries.

1. Tangential scanning is a true, volume CT scanner system which simultaneously collects data for an entire volume of an object. Data for many cross-sectional slices are simultaneously collected.
2. Tangential scanning has all the image quality advantages of a single-detector system because most tomographs are generated using data from a single detector.

3. Scanning speeds are even faster than third- and fourth-generation systems, because only the minimum amount of data is collected for an object of any size, and no time is lost in waiting for slice-to-slice movement or start/stop mechanics.
4. Data sets with any number of rays and views can be collected by changing the data collection rate, by adjusting rotation and translation speeds. This allows the system to achieve any desired geometrical resolution. The number of rays through the object is equal to the number of rotations during data collection. The number of rays through the object can be increased to achieve better spatial resolution, or they can be decreased to reduce scan time. Thus, the tangential system can collect data as if it has any (variable) number of detectors. Similarly, the number of views through the object is equal to the number of data points collected during a single rotation, i.e., the number of times per rotation that detector counts are recorded. Again, it can be increased for better spatial resolution or decreased for better scan time. Thus, the tangential system can collect data as if it has any number of views.
5. Extremely simple mechanical motions simplify the system's mechanical design and improve overall system reliability.

The only currently obvious limitations to tangential scanning are the unavailability of fast and effective reconstruction algorithms and the fixed pitch of cross-sectional tomographs (limited by detector width and spacing). Improved reconstruction algorithms are under development, however. Detector spacing can be fixed at a relatively small distance (currently 8 mm), and then particular detectors (every other detector, every third, etc.) can be read to obtain the desired pitch.

Scanner Prototype

An experimental apparatus has been designed and fabricated to collect data from logs up to 40 cm in diameter and 60 cm in length. This included a mechanical gantry with simultaneous translation and rotation of the log, a 128-channel detector array, a 300 kV X-ray generation system, fan beam X-ray collimation, and data collection, data analysis and image display software. A photograph of the apparatus appears in Fig. 6. More details of the apparatus and its operation can be found elsewhere [23].

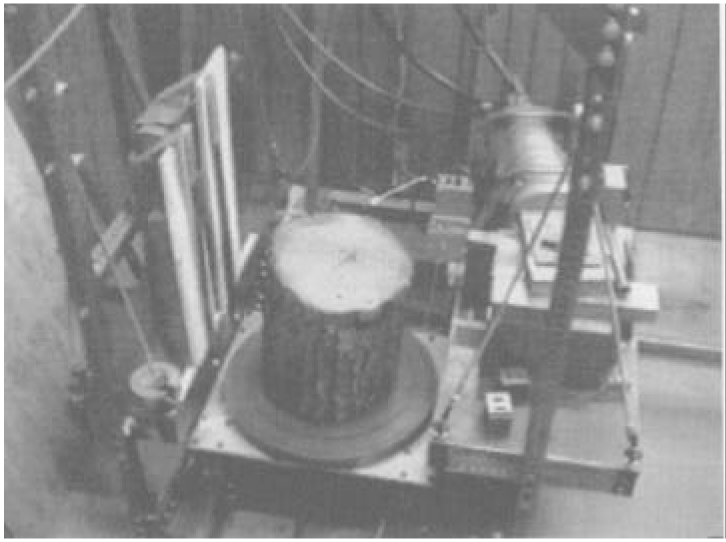

FIGURE 6 This photograph shows a bench prototype of a tangential scanner with a log section resting vertically on a turntable. which translates forward and back. A detector array is mounted vertically on the left, and the X-ray tube is mounted on the right.

In a typical experiment, the log rotates continuously with a rotational period of about 10 s per rotation. The linear period of the system is set to about 3200 s per m. Due to the belt-driven mechanics of the current system, there are significant variations in both the translational and rotational speeds. Nevertheless, this apparatus allows us to collect 1024 X-ray angular views per rotation and 3 rays per centimeter through the log during a typical tangential scan. Data collection is started manually and collects one line of data (128 readings – one from each detector) for each trigger pulse received from the encoder of the rotary motion. As the computer receives the data, it makes in-line offset and gain corrections on each reading for each detector before storing the data.

A typical tangential scan with 40 cm translation (360°) of the log through the X-ray beam produces approximately 32 MB of data. Other, larger data sets have also been collected using a slower motion of the translate stage, which generated more rays per centimeter. Filtered backprojection was performed on an individual detector to reconstruct

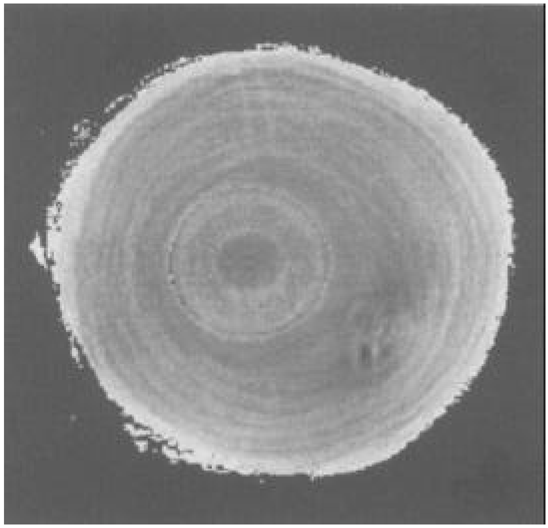

FIGURE 7 A CT reconstructed image of a log allows identification of heartwood and a knot.

an individual CT slice. An example CT reconstructed image from a slice of a softwood log appears in Fig. 7.

Despite demonstrated feasibility, the current prototype requires improvement before a full-scale prototype can be designed. First, direct drive systems for rotation and translation will contribute to improved speed and accuracy. Second, for a CT system used in log scanning, the current 1024 views is excessive and leads to more data than is really needed. We want to reduce the number of views to 600, which is an excellent number of angular views for a CT system with about 200 rays. Third, to counter the energy intensity drift of the current X-ray tube, it will be necessary to design, fabricate, and install a high-performance, low-noise single-channel X-ray detector, which can continuously measure the intensity of the X-ray tube. These reference detector measurements can then be used to correct the entire data set to eliminate the effect of the X-ray intensity drifts. Fourth, existing software needs to be extended to pre-process individual detector sinogram data, to filter sinograms to prepare for filtered backprojection, and improve backprojection of filtered sinogram data to reconstruct individual CT images.

The unique geometric design of tangential scanning provides cross-generational advantages without some of the inherent limitations, and permits collection of multiple slices simultaneously. These characteristics make it particularly effective for high-throughput, large-volume industrial inspection of hardwood logs. While this research program continues to extend tangential CT technology, we are also actively developing software to automatically locate and label internal features of logs from tomographs. This allows us to greatly condense the large amounts of CT data generated and to distill out the essential characteristics of logs to make processing decisions.

AUTOMATED INTERPRETATION OF CT IMAGERY

Generating CT images produces tremendous amounts of data. For example, depending on resolution and frequency of scans, the scan of a single 4-meter log may result in 20–800 MB or more of image data. Obviously, it is unrealistic to expect anyone to gain much insight into the 3-D appearance of an entire log by viewing a sequence of 2-D CT images. Fortunately, CT data contain a large amount of redundancy, which can be exploited to condense the data into a form that is more manageable and usable.

Only those internal features of a log that are important for subsequent processing need to be identified. These features are the defect areas within a log. Each density-related defect is relatively contiguous and each such defect type is fairly homogeneous with respect to density. Consequently, over the past 15 years researchers have begun to develop automated methods to interpret CT images [11,14,19,25–30]. Once different internal log defects can be automatically detected then it becomes a relatively straightforward task to integrate those views into a 3-D rendering of the log.

While previous efforts have demonstrated feasibility, serious limitations remain. First, reports of defect labeling accuracy are either anecdotal, based on success in a training set, or based on a single test set. Except for [25,27,28], no statistically valid estimates of labeling accuracy can be found in the literature. This makes it difficult to contrast the efficacy of competing approaches and to determine whether any particular approach can be used effectively in real-time scanning applications, Second, there has been no effort to assess or to achieve real-time operability of the developed algorithms. There seems to be a tacit assumption that computer hardware speed will eventually permit real-time execution of algorithms containing arbitrary complexity. Third, texture information (spatially contiguous and varying image elements), which is critical for human differentiation of regions in CT images (i.e., image segmentation), has been used for region labeling only [31].

Machine Vision Approach

The defect detection algorithm that we have developed [25,27,28] overcomes these three limitations. It consists of three parts: (1) a pre-processing module, (2) an artificial neural-net (ANN) based segmentation and classification module, and (3) a post-processing module. The pre-processing step separates wood from background (air) and internal voids, and normalizes density values. The segmentation-classifier labels each non-background pixel of a CT slice using histogram-normalized values from a $3 \times 3 \times 3$ or 5×5 window about the classified pixel. Morphological operations are performed during post-processing to remove spurious misclassifications.

Pre-processing

Background removal, which separates the wood region (foreground) from the background and internal voids, is the first objective of the pre-processing module. This step eliminates portions of the image from further analysis and, in turn, simplifies the classification procedure and decreases classification time. Background thresholding can be accomplished either statically or dynamically. This research applies Otsu's dynamic thresholding method [32]. Otsu's method works very well for bimodal histograms, but does poorly when histograms are multi-model. Because some log image histograms are multi-model, we have had to weight the histogram values before applying Ostu's method [28].

Normalizing CT image values is the second step of the pre-processing module. Because different species and different logs vary in density, somewhat different ranges of CT values can result. Histogram normalization translates the original CT image values into new values without disturbing the invariant associations that internal log features have with particular regions of the CT histogram. These associations seem to be, in our experience, consistent across many different species of logs in the green state (i.e., freshly cut).

ANN Segmentation/Classification

An ANN classifier is the core part of this classification system. Feed-forward back-propagation neural networks were chosen because their documented effectiveness for pattern-matching problems, and their relative ease of use. Using an ANN, each non-background pixel is labeled. We have constructed both 2-D (5 × 5) and 3-D (3 × 3 × 3) ANN classifiers, each with a single hidden layer and with a l-of-N output layer containing log feature types: clear wood, bark, voids, knots, decay and splits. Input layers contain the normalized pixel values for the target pixel’s local neighborhood (either 27 or 25 elements, one per neighborhood pixel) plus one additional element that contains the radial distance of the target pixel to the center of the log (Fig. 8). One of our primary research objectives was to determine if local texture information (augmented with some contextual information, radial distance) could be used to classify images.

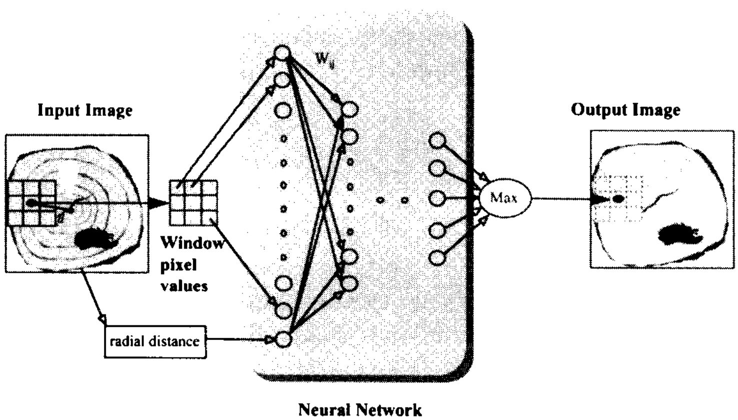

FIGURE 8 The layout of our artificial neural network classifier depicts the source of input nodes, the hidden layer and classifier output.

Based on prior results [25], a single hidden layer containing 12 nodes is used. The numbers of output nodes for the ANNs differ, however. For example, red oak classifiers detect five classes: clear wood, knots, bark, splits and decay. Our yellow poplar samples, on the other hand do not contain decay, but do contain sapwood, which is different in density than heartwood for this species, and yet both (sapwood and heartwood) are clear wood components. In 2-D classifiers, the topology is 26-12-5 or 26-12-6, which means that the structure of the neural network has 26 input nodes, 12 hidden nodes, and 5 or 6 output nodes. In 3-D classifiers, the topology is 28-12-5 or 28-12-6, which has a similar interpretation.

Post-processing

Because classification features are based primarily on local neighborhoods, spurious misclassifications tend to occur at isolated points. A post-processing module is used to remove these small regions, and therefore improve overall system performance. After passing an image through an ANN classifier, a CT image is labeled and treated as a gray-level image. Then the image is post-processed by the morphological operations of erosion followed by dilation using a 5-point structuring element. Splits are delicate features and, if post-processed, are often deleted by the erosion operation. Hence, for all classifiers in our study, an entire image is not post-processed, but only the outer regions of the log, because splits tend to lie near the log center. This approach deletes misclassified small areas – which occur mostly near the outer edges of the log – and yet retains important information (like splits) near the center of the log.

Defect Recognition Accuracy

An entire training/testing set for one hardwood species consists of approximately 1000 samples across multiple images. Ten-fold cross validation was used to evaluate the accuracy of each classifier. This means that a training set is randomly divided into 10 mutually exclusive test partitions of approximately equal size. For each of the 10 stages of training, one partition is designated as the test set, and the remaining samples in other partitions are used to train the neural network. In successive stages, different partitions are used for testing and the remaining samples are used for training. The average classification accuracy over all 10 stages of training is reported as the cross-validated classification accuracy.

All the ANNs were trained using the delta rule. Based on Li's results [25], a small learning rate 0.1 and a medium momentum term 0.6 were selected as the learning parameters for all ANNs. Random values were assigned to the initial weights for each network training session.

Using 10-fold cross-validation we developed individual classifiers for each species – red oak, yellow poplar, and cherry – using both 2-D and 3-D feature vectors (6 classifiers). Image pixels were nominally $(2.5\ \text{mm})^3$ resolution. We also developed multiple-species classifiers: pairing two species at a time and combining all three species together. These were also trained using 2-D and 3-D feature vectors for a total of eight multiple-species classifiers. Finally, finer resolution cherry images $(0.95\ \text{mm})^3$ were used to train both a 2-D and 3-D classifier. Classification accuracies appear in Fig. 9.

The accuracy of all six single-species classifiers is above 95%. Six, two-species classifiers have also been trained using both 2-D and 3-D image data. Their accuracy is 90–97%. Finally, combined three-species classifiers (red oak, yellow poplar and cherry) were generated for 2-D and 3-D analysis. These two classifiers identified six kinds of

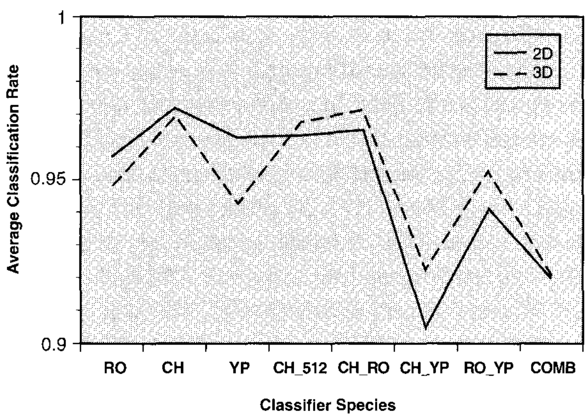

FIGURE 9 2-D and 3-D classifier accuracies are plotted for each of the ANN classifiers – red oak (RO), cherry (CH), yellow poplar (YP), 512 × 512 cherry (CH_512) cherry/red oak (CH_RO), cherry/yellow poplar (CH_YP), red oak/yellow poplar (RO_YP) and all 3 species combined (COMB).

defects: clear wood, knot, bark, split, decay and yellow-poplar sapwood. Their accuracy is about 91–92%. All of these classification accuracies are prior to post-processing. Visual assessments indicate that post-processing operations improve accuracy even further Fig. 10.

Defect Labeling Generality

By training classifiers with different species and with different pixel neighborhoods, we were able to investigate the interaction of neighborhood shape (2-D vs. 3-D) and single- vs. multiple-species classifiers, with respect to their impact on classifier accuracy. The issue that we sought to resolve here is whether we could develop species-independent classifiers of high accuracy using our ANN, local-neighborhood approach.

Therefore, the results in Fig. 9 were examined statistically. This type of analysis is possible because each estimate of classification accuracy is an average of 10 sample estimates for the individual cross-validation partitions. We used Analysis of Variance along with post-hoc T-tests to answer questions on the impact of neighborhood shape (2-D vs. 3-D) and classifier cardinality (single- vs. multiple-species) on classification accuracy. In our first statistical test, we found a significant interaction between shape and cardinality. This interaction can be seen in the average classification rates of Fig. 9, where 2-D rates are generally higher for single-species classifiers and 3-D rates are generally higher for multiple-species classifiers. In a second statistical test, we found that differences existed among the set of single-species classification rates, and also among the set of multiple-species classification rates. In both cases, the rates for 2-D and 3-D neighborhood differed statistically.

In a final test, we excluded classifiers based on *both* cherry *and* yellow poplar data and performed our original ANOVA again. The fine resolution (0.95 mm)' cherry classifier (CH_512) was also excluded. As before, we blocked the ANOVA on shape (2-D and 3-D). The resulting F-ratio value for cardinality indicates that there is no difference between single- and multiple-species classification rates when cherry/yellow poplar combinations are removed.

We have formed two significant generalizations from these results. First, when comparing single-species classifiers and multiple-species

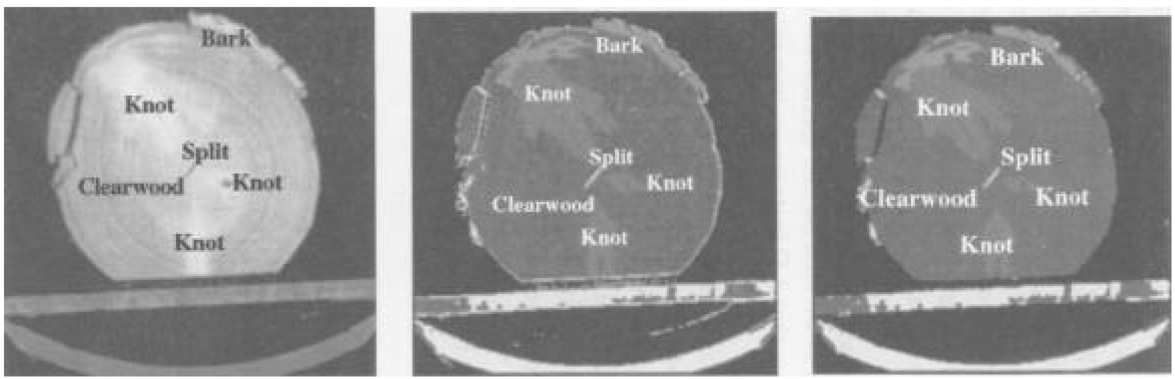

FIGURE 10 A red oak classifier is applied to a red oak image (left) producing the classified image (center), which is finally post-processing (right).

classifiers, the performance of the former is better than that of the latter when cherry–yellow poplar combinations are used. On the other hand, when those combinations are excluded, there is no significant difference between classification accuracy for single- and multiple-species classifiers. Second, when comparing 2-D and 3-D features, the performance of 2-D single-species classifiers is better than that of 3-D single-species classifiers. The performance of 3-D multiple-species classifiers is better than that of 2-D multiple-species classifiers. We conjecture that for accurate classification in single-species classification, multiple-image planes contain redundant data that may be unimportant, or even counter-productive. For multiple-species classification, however, the extra information contained in adjacent CT slices seems to aid feature labeling. Consequently, as we increase the species mix that a classifier must deal with, it appears that 3-D features are important for attaining high accuracy.

This segmentation/classification technique is able to label an entire CT image (containing 64K pixels) in l–2 s. In addition, our results to date indicate that it can be trained to work for any species with a statistically valid accuracy rate of 95–98% at the pixel level [25,28]. After this defect detection algorithm is applied to each CT image for a log, slice-by-slice data regarding each defect and the log perimeter can then be used to generate “glass log” images. These images can be viewed by a sawyer prior to log breakdown or can be used to evaluate alternative sawing patterns in software.

APPLICATION OF NDE INFORMATION

The current work of improving yield from hardwood processing using an integrated approach is predicated on the availability of internal imaging information. With such information available, it is feasible with current technology to model the log through all of its processing steps – not just log breakdown. A internal log model can be fed back and incorporated into decision-making during log breakdown, edging and trimming, grading and sorting, drying, cross-cutting and ripping, matching and gluing, and eventual end-use manufacturing [33]. The log model used is a solid model that supports Boolean operations to mimic various processing operations [34].

Currently, the application of NDE information is focused on the primary processing activities occurring in the sawmill, including log breakdown and board edging and trimming. However, in the not-too-distant future, we foresee this work being extended to the downstream operations of cross-cutting and ripping in the roughmill to produce dimension parts. It also appears that this idea could be extended to earlier log processing stages, log bucking and topping decisions (Fig. 1). Due to the inherently interrelated nature of hardwood log processing, all of these manufacturing steps can be integrated for optimal results. The following sections describe some of the work that has resulted from the application of NDE information in hardwood log processing.

Data Reduction

One of the constraints in modeling, rendering, and processing information derived from NDE imaging is dealing with huge data sets. Even after background removal and labeling, the data remain large, approximately 8–10 MB per log. In a previous study, for example, red oak logs measuring 10–12 feet (2.5–3 m) in length were scanned every quarter-inch (6.35 mm) to detect the occurrence of internal defect information [35]. The significant cross-sectional changes in a log profile, however, do not occur in small increments of several millimeters, but rather over a range of several centimeters. Likewise, each cross-sectional log profile contains more data points than may be necessary to adequately describe its shape. To represent a log profile then, it is possible to distill the significant data and not have to carry the overhead of a massive data set. For defect profile representation, the same applies, but at the quarter-inch scale.

To speed up processing and better manage the data, it is desirable to reduce the data to a minimum set that still retains critical shape information. A computer model called GDR (for Geometric Data Reduction) has been developed that reliably reduces a log's data set to a more manageable size (~ 600 KB) while preserving the representational integrity of the geometric information. The model essentially eliminates slice data (in a recursive fashion) which do not exhibit unique centroidal displacement or size characteristics within a threshold value. Figure 11 shows a comparison of a log before and after processing by

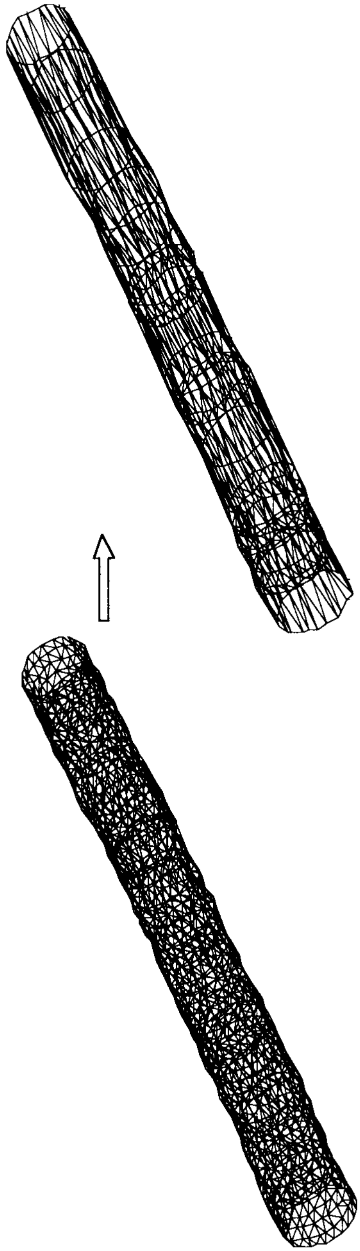

FIGURE 11 A log before (left) and after (right) geometric data reduction illustrates a substantial reduction in log slices.

GDR. The reduction ratio for the figure is 0.142, which is equivalent to a 59% reduction in the number of slices for only a 0.03% reduction in solid volume [36].

Log Sawing Simulation

After data reduction, the geometric information is converted into a polyhedral solid model of the log and its defects [37]. A polyhedral solid model not only approximates the true shape of the log and its defects more closely than previous models (e.g. [38]), but also provides a more robust model that includes both geometric and topological properties, and enables manipulation through regularized Boolean operations.

Introduction of a polyhedral solid modeling approach for hardwood log and defect representation facilitated the development of an interactive graphics-based sawing simulator. The sawing simulation program called GRASP (for GRAphic Sawing Program) is a microcomputer-based graphics program that enables simulated sawing of solid representations of a log with embedded defects [34].

GRASP has all the attributes of a computer-aided design (CAD) graphics modeler, such as window viewing, hidden line redering, spatial transformations, and geometric calculations, making it a powerful tool for sawing simulation. It is flexible enough to use in any sawing operation, from log bucking, topping, log breakdown, quartering, veneering, to edging, trimming, secondary processing, even extracting and representing furniture components.

As an NDE application, the simulator can be used to view the log both as an opaque, as well as a transparent, container for defects, simply by hidden-line removal rendering. Figure 12 shows a quartersawn solid log representation.

Log Processing Integration

The integrated approach we are taking is grounded in the fundamental philosophy of integration. The basic premise of integration is that some decisions are interrelated, and thus these decisions should not be made separately in isolation [5]. Examples of potential candidates for integration can be found at different stages of hardwood processing [33].

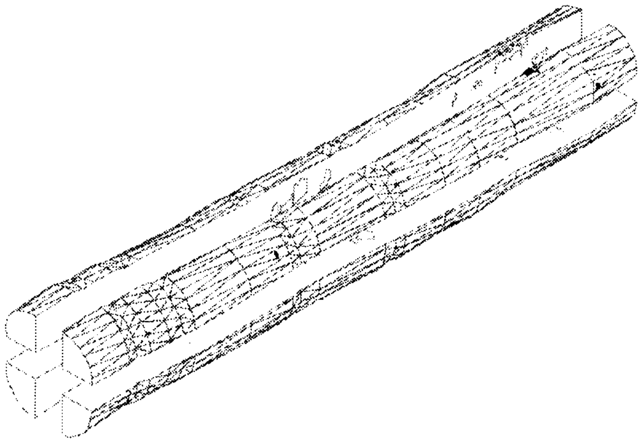

FIGURE 12 Hidden-line removal rendering produces the typical view of a log solid.

Take, for example, the interaction between the sawyer and the edger operator. When the sawyer removes a flitch (a board containing wane, i.e. bark on its edges) from a log face, there is an expectation of a potential grade for that face. This expectation, however, is not communicated as information to the edger operator who may remove too much or too little wane, resulting in a different board grade from that which the sawyer intended.

Another example, on a different scale, is an end-use manufacturer who may need a 3-inch (7.62 cm) thick piece of wood as a furniture component. To arrive at this dimension from a batch of l-inch (2.54 cm) lumber or dimension stock, the manufacturer has to match and glue together several pieces. From an integrated viewpoint, this dimension requirement can be communicated as information to the sawyer who will then saw a 3-inch (7.62cm) thick flitch for this manufacturer and save a few intermediate steps.

The hardwood industry is quite segmented, with each segment usually concerned only with its immediate input (supplier) and output (customer). As such, a sawmill converts logs to lumber for the secondary market, which in turn converts lumber to dimension stock, then ultimately to the furniture manufacturer, panel manufacturer, and other end users. This segmentation is deeply rooted in the current market structure. Our integrated approach runs counter to this market segmentation, capitalizing on inherent product interrelationships to develop products at a lower cost with less waste. Earlier work directed at examining direct log-to-dimension manufacturing [39], and more recently, work on direct log-to-furniture manufacturing [40] are examples of integrated approaches.

NDE facilitates the integration of hardwood processing by providing a common basis for decision-making. When internal defects are exposed at the very start of manufacturing, it gives downstream processes an entirely new spin. Knowing the type of internal defects and their distribution and orientation inside a log significantly affect the sawing pattern, expected lumber sizes, intermediate processing steps, and ultimately the grading valuation. With log NDE, more information becomes available. The challenge is how to effectively use the information to one's advantage.

Log Breakdown Analysis

Traditionally, log breakdown follows a few sawing patterns: live sawing, grade or around sawing, cant sawing, taper sawing. Usually a sawmill adopts one of these sawing patterns and uses it consistently on most of its logs. With additional information available through NDE, regarding internal defect configurations, it is conceivable for logs with different defect configurations to be subjected to different sawing patterns on a case by case basis [33]. One early computer model designed to deal with this defect-specific approach was PDIM (Pattern Directed Inference Model) which generates a log breakdown pattern specific to the internal defect configuration found inside the log [4l]. It accomplishes this by enveloping the defects in a defect hull and analyzing a composite end-view that represents an aggregation of the defects' distribution through the log. The automated decision-making was driven by the shape of the hull, and density numbers that reflected the defect concentration along the length of the log. Designed to be a generative process planning model, PDIM generates sawing instructions that could be used to direct a numerically-controlled sawing headrig and log carriage. This model, along with other similar models, is being investigated for effectiveness in arriving at computer-generated optimal sawing patterns.

In a pilot study involving three log grades of varying quality (grades #1, #2 and #3, with #1 being the highest quality), six sawing heuristics were applied in both, defect information-limited (traditional) and information-augmented (with NDE) scenarios [42]. The six heuristics were obtained from popular sawing practice as reported by Malcolm [43]. Preliminary results indicate that in the absence of an optimal log breakdown procedure, increased information about internal log features can improve value recovery by 8.5% for grade #1 logs. Lumber values for lower grades did not change significantly, which suggests that choosing a breakdown pattern with high recovery becomes very difficult when viewing logs with many defects. A follow-up study involving a larger sample of logs is underway.

CONCLUSIONS AND DISCUSSION

The application of NDE methodology to hardwood log sawing is a challenging research problem for several reasons. First, tremendous amounts of data are generated for each unit (log) that is imaged. Therefore, an important focus of our image-analysis work has been data reduction – condensing CT imagery data down to defect and log profiles. Further, our approach also reduces the number of profiles that must be processed during simulated sawing and 3-D rendering. Second, the eventual application of this technology requires real-time, in-line NDE and data processing. Sawmill profit margins are relatively small and cannot support additional log handling and sorting required for off-line operations. Third, wood possesses tremendous internal heterogeneity and biological variability, both across and within species. Data processing software must be robust and dynamic to deal with this variability. Finally, the technology being developed is at this stage both complex and expensive, which runs counter to the generally less sophisticated and lower-capitalized, hardwood processing industry.

Nevertheless, significant progress has occurred in all three areas described above: scanning technology, image analysis and data utilization. Work continues on the tangential scanning bench prototype and on automated defect detection, which will be extended to additional species and enhanced by better post-processing methods. A full-scale prototype is contingent on a future support from the private sector.

In the meantime, simulated log sawing training software is currently being developed with industry support. It utilizes the software developed previously for data reduction, log sawing and log breakdown analysis. The trainer will allow sawyers to practice sawing logs in a realistic environment with feedback on performance. Presently, sawyers destructively improve their sawing skills using a mill's inventory, and profits.

Current sawing heuristics applied by sawyers are based on externally visible log characters and those defects exposed during sawing. These heuristics eventually need to evolve as internal scanning becomes operational. We are studying the impact that viewing a rendered 3-D glass log will have on those traditional heuristics, and additionally, we are developing new heuristics to incorporate internal information [42].

The ultimate goal of this research program is to develop the scientific and technological foundation that is needed to make operational NDE practical for sawmills. Hardwood inventories (raw materials) and the sawmilling business environment are changing. Raw material availability and quality are decreasing and cost is increasing, while lumber prices (salable product) are relatively stable. The resulting small profit margins in the hardwood sawmilling industry mean that even small increments in value recovery translate into large percentage increases in profits. While the technology being developed is not inexpensive, it will become necessary for sawmills of the future.

References

[I] R.J. Bush, P.A. Araman and J. Muench, Jr., in *Wood Product Demand and the Environment,* Forest Products Society, Madison WI (1992).

[2] D.B. Richards, W.K. Adkins, H. Hallock *et al.,* Forest Products Lab, Madison WI, Res. Pap. FPL-355 (1979).

[3] J.A. Tsolokides. *Forest Products Journal,* **19,** 7 (1969).

[4] F.G. Wagner, T.E.G. Harless, P.H. Steele *et al.*, in *Proceedings of Process Control/ Production Management of Wood Products: Technology for the 90's,* Athens GA, The University of Georgia, Athens GA (1990).

[5] L.C. Occeña, *Journal of Forest Engineering,* **3,** 1 (1991).

[6] S.J. Chang, P.C. Wang and J.R. Olson. in *2nd International Conference on Scanning Technology is Sawmilling,* Oakland/Berkeley Hills CA, R. Szymani (Ed.), Forest Industries/World Wood, San Francisco CA (1987).

[7] W. Han and R. Birkeland, *Industrial Metrology,* **2,** 3/4 (1992).

[8] D.M. Benson-Cooper, R.L. Knowles, F.J. Thompson *et al.*, Forest Research Institute, New Zealand Forest Service. Rotorua NZ, Bull. No. 8 (1982).

[9] A.E. Burgess, in *1st International Conference on Scanning Technology in Sawmilling,* San Francisco CA, R. Szymani (Ed.), Forest Industries/World Wood, San Francisco CA (1985).

[10] D.J. Cown and B.C. Clement, *Wood Science and Technology*. **17,** 2 (1983).
[11] C.W. McMillin, *Wood Science,* **14,** 3 (1982).
[12] M. Onoe. J.W. Tsao, H. Yamada *et al., Nuclear Instruments and Methods in Physics Research,* **221,** 1 (1984).
[13] F.W. Taylor, J.F.G. Wagner, C.W. McMillin *et al., Forest Products Journal,* **34,** 5 (1984).
[14] B.V. Funt and E.C. Bryant. *Forest Products Journal,* **37,** 1 (1987).
[15] P.O.G. Hagman and S.A. Grundberg, *Hols als Roh- und Werkstoff,* **53** (1995).
[16] S. Som, P. Wells and J. Davis, in *Second International Conference on Automation, Robotics and Computer Vision,* Singapore (1992).
[17] D.P. Zhu, R.W. Conners and P. Araman, in *Proceedings of the 23rd Southeast Symposium on System Theory,* Columbia SC (1991).
[18] P.A. Shadbolt, Department of Applied Physics, Chisholm Institute of Technology, M.S. Thesis (1988).
[19] F. Hopkins, I.L. Morgan, H. Ellinger *et al., Materials Evaluation,* **40,** 20 (1982).
[20] S.J. Chang, in *3rd International Conference on Scanning Technology in Sawmilling,* San Francisco CA, R. Szymani (Ed.), Forest Industries/World Wood, San Francisco CA (1989).
[2l] D.G. Hodges, W.C. Anderson and C.W. McMillin. *Forest Products Journal,* **40,** 3 (1990).
[22] D.L. Schmoldt, in *Proceedings of the 1996 Hardwood Research Symposium,* D. Meyer (Ed.), National Hardwood Lumber Association, Memphis TN (1996).
[23] N.K. Gupta, D.L. Schmoldt and B. Isaacson, in *Multisource-Multisensor Information Fusion,* Las Vegas NV, H.R. Arabnia and D.P. Zhu (Eds.), CSREA Press (1998).
[24] N.K. Gupta, Tangential computerized tomographic scanner, Omega International Technology, Inc., Patent No. 5,648,996 (1997).
[25] P. Li, A.L. Abbott and D.L. Schmoldt, in *Proceedings of the 1996 IEEE International Conference on Neural Networks,* Institute for Electrical and Electronics Engineers, Inc., Piscataway NJ (1996).
[26] D.L. Schmoldt, D.P. Zhu and R.W. Conners, in *Review of Progress in Quantitative Nondestructive Evaluation,* D.O. Thompson and D.E. Chimenti (Eds.), Plenum Press, New York (1993).
[27] D.L. Schmoldt, P. Li and A.L. Abbott, *Computers and Electronics in Agriculture,* **16,** 3 (1997).
[28] D.L. Schmoldt. J. He and A.L. Abbott, in *Photonics West 1998,* San Jose CA, SPIE, **3306** (1998).
[29] D. Zhu, R.W. Conners, F.M. Lamb *et al.,* in *4th International Conference of Scanning Technology in Sawmilling,* San Francisco CA, R. Szymani (Ed.), Forest Industries/World Wood, San Francisco CA (1991).
[30] D. Zhu, R.W. Conners, D.L. Schmoldt *et al., IEEE Transactions on Systems, Man, and Cybernetics,* **26,** 8 (1996).
[31] D. Zhu and A.A. Beex, *Journal of Visual Communication and Image Representation,* **5,** 1 (1994).
[32] N. Otsu, *IEEE Transactions on Systems, Man and Cybernetics,* **SMC-9** (1979).
[33] L.C. Occeña, D.L. Schmoldt and P.A. Araman, in *Twenty-Fourth Annual Hardwood Symposium,* Cashiers NC, D. Meyer (Ed.), National Hardwood Lumber Association, Memphis TN (1996).
[34] L.C. Occeña and D.L. Schmoldt, *Forest Products Journal,* **46,** 11/12 (1996).
[35] D. Zhu, R.W. Conners, D.L. Schmoldt *et al.,* in *1991 IEEE International Conference on Systems, Man, and Cybernetics,* Charlottesville VA (1991).
[36] L.G. Occeña, W. Chen and D.L. Schmoldt, in *4th Industrial Engineering Research Conference,* Nashville T.N., B. Schmeiser (Ed.) (1995).
[37] L.C. Occeña and J.M.A. Tanchoco, *Forest Products Journal,* **38,** 10 (1998).
[38] S.M. Pnevmaticos, P.E. Dress and F.R. Stocker, *Forest Products Journal,* **24,** 3 (1974).

[39] P.A. Araman, W. Lin and D.E. Kline, in *Twenty-Third Annual Hardwood Symposium,* Cashiers NC, G. Lowry and D. Meyer (Eds.), National Hardwood Lumber Association. Memphis TN (1995).
[40] L.H. Wang and L.C. Occeña, Department of Industrial Engineering, University of Missouri-Columbia, Technical Report #WR980701 (1998).
[41] L.G. Occeña, *Wood and Fiber Science,* **24,** 2 (1992).
[42] L.G. Occeña, D.L. Schmoldt and S. Thawornwong. in *ScanPro: Advanced Technology for Sawmilling,* J. Dennig (Ed.), Miller-Freeman, San Francisco (1997).
[43] F.B. Malcolm, Forest Products Lab, Madison WI, Report No. 2221 (1961).