# Automated detection of severe surface defects on barked hardwood logs

Liya Thomas
Clifford A. Shaffer
Lamine Mili
Ed Thomas

**Abstract**

We have developed an automated detection algorithm that identifies severe external defects on the surfaces of barked hardwood logs and stems. To summarize the main defect features and to build our defect knowledge base, we measured, photographed, and categorized hundreds of real log defect samples. Three-dimensional laser-scanned range data captured the external log shapes and showed bark pattern, defective knobs, and depressions. Severe defects were identified via the analysis of 3-D log data using decision rules derived from analysis of the knowledge base. Defects were detected by examining contour curves generated from radial distances determined by robust 2-D circle fitting to the log-data cross sections. There were a total of 68 severe defects, of which 63 were correctly identified. There were 10 nondefective regions falsely identified as defects.

Automatically locating and classifying log defects helps to improve lumber yield, in terms of both volume and quality. Traditional defect inspection consists of a very brief visual inspection by the sawyer. Visual inspection has a high error rate and is easily influenced by the operator's physical and mental conditions. Thus, researchers have been developing a variety of computerized defect detection and classification systems to assist the sawyers' decision-making process (Chang 1992).

Computerized tomographic/x-ray technology has been used to locate internal hardwood log defects in the laboratory (Zhu et al. 1991, Li et al. 1996). Log defects exist both externally and internally. As CT/x-ray technology is capable of penetrating material, the resulting images display internal defects through density variations. While CT/x-ray-based detection approaches generate successful experimental results with a 95 percent detection accuracy (Li et al. 1996), several obstacles have prevented them from being used in industrial applications. First, the data collection speed is extremely slow because of the large data volume, varying anywhere from 5 minutes to 4 hours per log. Second, variation in MC in the log causes the intensity of scanned images to vary, making detection results unstable. Third, the technology presents an environmental hazard, as penetrating such a large object requires a tremendous amount of x-ray energy. Finally, the high cost of the scanning equipment—on average 1 million, or more, U.S. dollars—is beyond most sawmills' reach and thus has little practical value.

In contrast, 3-D laser scanner technology uses relatively low-cost equipment (US$250,000 to US$400,000 ) that is more affordable to sawmills. Laser scanning equipment collects the external log shape information using triangulation technology. Since only surface data are collected, data collection speed is much faster and the equipment can operate at line speed (seconds per log). The system employs low-energy laser-scanning units, which are safe to operate. MC does not interfere with 3-D profile data.

The laser-scanning system used in our research is a commonly available industrial system manufactured by USNR, Inc.[1] The scanner generates high-resolution profile images of

[1] The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the U.S. Department of Agriculture or the Forest Service of any product or service to the exclusion of others that may be suitable.

The authors are, respectively, Post-doctoral Research Associate, Associate Professor, Computer Sci. Dept., Virginia Tech, Blacksburg, Virginia (lithomas@vt.edu, shaffer@vt.edu); Professor, Electrical Engineering Dept., Advanced Research Inst., Virginia Tech Northern Campus, Arlington, Virginia (lmili@vt.edu); and Research Scientist, USDA Forest Serv., Princeton, West Virginia (ethomas@fs.fed.us). This paper was received for publication in February 2006. Article No. 10171.

Forest Prod. J. 57(4):50-56.

the log surface in three dimensions. The scanner was primarily developed for the softwood industry, where the scanner is used to determine the shape and size of the log being sawn in three dimensions. Ideally, an optimizer would take the scan data and determine the optimal sawing pattern for the log. The system's resolution is high enough such that defects can be manually located in the scan data by the human eye. The obvious question then is how to get the computer to see the defects too.

Most severe log defects are associated with a localized, significant height rise. Severe defects are those that significantly degrade the value and/or strength of the boards that would be sawn from the log (Carpenter et al. 1989). To detect severe defects, we developed an automated defect detection algorithm using laser-scan profile data. To begin, circles were fit to data cross sections and the radial distances between the fitted circle and the data calculated (Thomas and Mili 2006). From the radial distances we generated a grayscale image showing the height changes of the log surface. This image is then used to determine a contour plot of the log surface from which the large and/or protruding defects are determined. However, some types of severe defects do not present significant height change against the surrounding bark, and thus are not detected by the algorithm presented in Section 3. We hope to develop pattern-based methods to identify these kinds of defects in future work. For this paper, we examine only those defects with a significant height rise.

The main disadvantage of the laser scan method is that it only provides external defect information, which might prove insufficient for lumber processing. To address this problem, a sister study (Thomas et al. 2006) to determine the correlation of external and internal defects is ongoing at our partial sponsor, the USDA Northern Research Station in Princeton, West Virginia. Strong correlations have been found to exist between external indicators and internal characteristics. For the most severe defects, the models can predict internal features such as total depth, midway point defect width and length, and penetration angle, with a low measurement error. For less severe defects such as adventitious knots and medium and light distortions, the correlations are somewhat weaker. To the best of our knowledge, we are the first group investigating detection methods for defects on the surface of hardwood logs and stems using laser-scanned 3-D Cartesian coordinates (Thomas et al. 2003, Thomas et al. 2004).

## Objectives and methods

We obtained log data from two commercially important northeastern American hardwood species: yellow-poplar and red oak. More than 160 log data samples were collected, each consisting of a series of cross sections along the log length at 0.8-inch intervals (**Fig. 1**). Each cross section comprises approximately 1,000 3-D coordinates with adjacent points roughly 0.05 inches apart, so the data are much denser around the logs circumference than along the length. Typically a log's length ranges between 8 and 16 feet. Thus, one log data sample has between 120,000 to 240,000 points.

The frame of the log scanner blocked the laser signal wherever the log was supported. This is the cause of the white bands shown in **Figure 1**, where 5 to 7 data circles are incomplete for those two places. In addition, the frame produced severe outlier issues on all logs. Calibration problems with the scanning units and log diameters also caused missing or duplicated data. The nature of the log data, with the large overall quantity and small percentage of severe outliers, calls for modeling the data using robust curve-fitting methods rather than conventional least-squares fitting. This leads us to the application of robust statistics and the development of our 2-D circle-fitting Generalized-M Estimator (GME) (Hampel et al. 1986, Thomas et al. 2003, Thomas and Mili 2006).

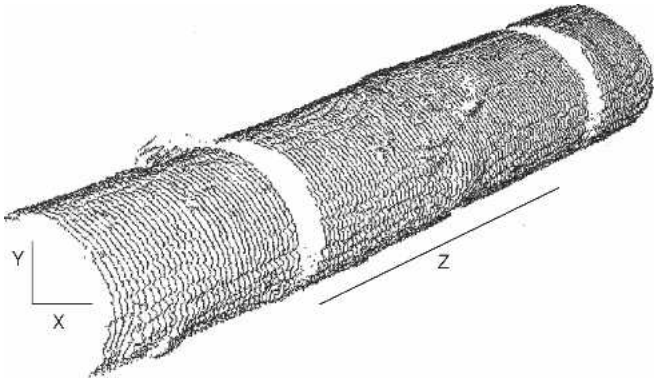


*Figure 1. — Dot cloud projection of 3-D log data. Shown is part of the data for one log sample. A bump-like external defect (lower left), missing data, and outliers caused by loose bark (upper-middle left) are visible.*

Actual defect locations, sizes, types, etc. were measured manually on the log samples. Color digital images of the log surface, four images per log (at 90° intervals), were taken as well. About 500 samples of external defects were studied, measured, and photographed. These defect samples were analyzed to provide indicators and classification of external defect characteristics. Statistics for these defect classifications are used to define our defect-detection algorithm and to improve it through comparing its simulated data generated by the algorithm against the actual defect statistics.

## Detection algorithm

The procedure for detecting external defects consists of two major steps. The first step is to obtain the radial distances by fitting 2-D circles to log-data cross sections using a robust GM-Estimator (Thomas et al. 2003). The circle-fitting algorithm is written in Java. Original log data points are processed by removing outliers using the fitted circles with a threshold of 2 inches for their radial distances. Then the data points (x and y coordinates and radial distance) are sorted according to the angles of vectors passing through the circle center and points. The approach described here requires that there be no missing data. Thus the algorithm "fixes" regions with missing data in the matrix of radial distances by using a linear interpolation. The final product of this step is a grayscale image with pixel values indicating radial distances from the fitted circles to the actual log data (see **Figure 2**). The second step is to determine the locations of severe defects on the log surface. Our current implementation for this phase is in Matlab 7. The detection program incorporates expertise obtained through measuring, photographing, and analyzing approximately 500 external-defect samples.

Before describing our detection algorithm, we must first define the "defects" we are looking for. The resolution of the laser scanner (0.8 inch along log length) restricted the size of defects that can be reliably detected to 5 inches and greater in diameter. Our current detection algorithm only detects defects

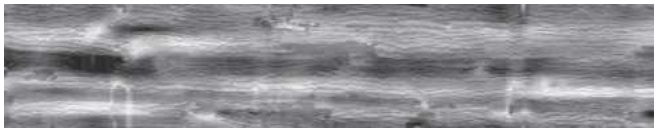

*Figure 2. — A grayscale image of radial distances of a log sample. The bright regions illustrate large radial distances, and thus indicate protruding defects. The dark areas show small distances with respect to the reference fitted circles, and thus might indicate defects such as holes, splits, and gouges. This image reflects height changes on the log surface.*

with a minimum 1-inch surface rise, because it is height (surface rise) based. Thus, we define "severe defects" to mean those with at least 1-inch surface rise, 5 inches in diameter, and a width to length ratio between 0.5 and 2. From the 160 logs scanned, 14 representative samples were selected. From these samples we observed a total of 60 severe defects. An additional 8 less severe defects were observed on the 14 logs. Less severe defects are those without significant height change (0.5 to 1 inch) and have a distinctive bark pattern and a diameter of 3 to 5 inches. Below is an overview of the detection process in pseudo-code:

1. Find severely protruding (≥1 inch in height) and large (≥5 inches in diameter) defects. Using radial distance data, obtain contours at six evenly spaced levels from radial distances. A contour or contour curve in a plot is a curved line connecting points with the same surface rise. The first contour level is the lowest, and sixth is the highest. The algorithm retains only level 6 contours. From this point, most processing is on the bounding boxes (regions) (**Fig. 3**).
   a) Eliminate regions whose area is less than 5 $in^2$. A rectangular region, or simply a region, is a solid rectangle enclosed by the bounding box for a contour.
   b) Sort regions in descending order of area.
   c) Eliminate long and narrow regions as these do not correspond to typically defective areas.
   d) Adjust bounding boxes for contours by determining whether each box encloses entire sawn tops of branch stubs; we refer to these as adjusted regions. Remove adjusted regions with severe missing data, and remove adjusted regions that are too small.
   e) The remaining regions are reported as possible defects.
2. Find the less protruding (≤1 inch in height) and smaller (≤5 inches in diameter) defects:
   a) Using the original 3-D log data, determine gradients parallel to the long axis of the log.
   b) Find the areas with gradients within the defined range for this defect class. These areas are reported as defects.

A built-in Matlab function converts the grayscale image to a contour plot. The function inputs and analyzes radial distances generated by the circle-fitting procedure to locate areas where surface defects might exist. First, it obtains the contour curves based on the radial-distance data. The original 3-D log data are then read in. Depending on the scanner calibration and the diameter of the log, the original log data may contain some replicated points. The algorithm removes the replicates. For each data point, a line is drawn from the point to the cross

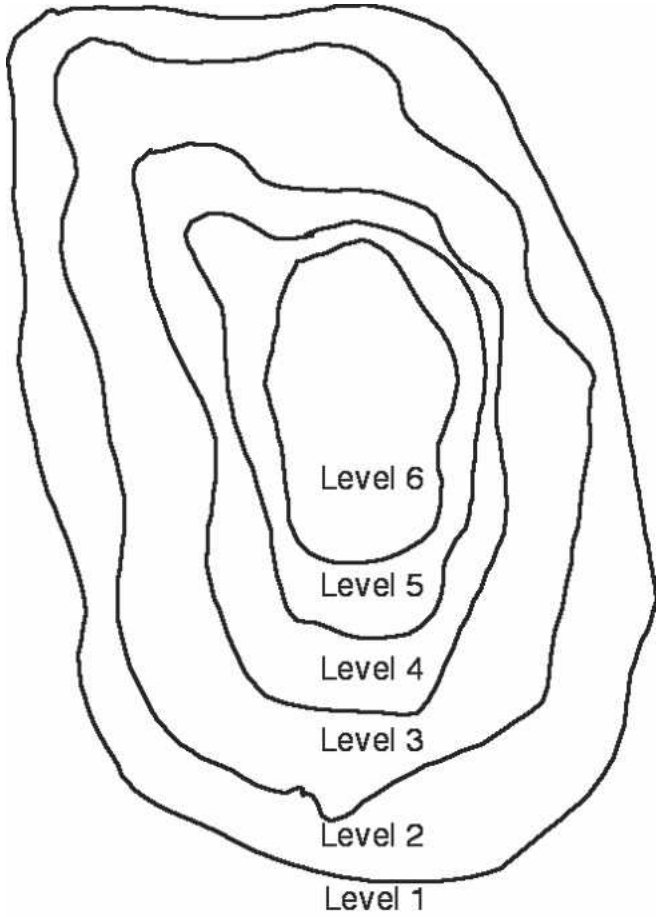


*Figure 3. — Illustration of contour curves at 6 different levels. Currently, we process only the sixth, or highest, level.*

section's fitted circle center. The angle between this line and a horizontal line is computed. The points on a cross section are then sorted by their angle values. Second, for each contour curve, the algorithm determines its borders. The width, length, area, width/length ratio, and length/width ratio are computed. Presently, we only analyze the highest (sixth) level contours, which enclose the highest regions and the most protruding defects. Typically, each log has anywhere from a few dozen to a few hundred contour curves at the highest level.

The key objective in the remainder of the algorithm is to identify possible defect regions through a series of steps to eliminate non-defective regions from the potential candidates. This is achieved by using statistics from measured and calculated log data and log feature expertise in a stepwise fashion. The algorithm removes the regions whose area is less than 5 $in^2$ because the data resolution (0.8 inch between cross sections) means the regions cannot be recognized as defects. Next, we sort the remaining regions in order of their areas. We do this so that it is efficient to determine whether a smaller region is nested inside a bigger one. Any contour nested within another is removed from consideration because there can only be one defect in a given location.

To obtain an estimation of potential defect locations at the beginning of the algorithm, only the widths and lengths of contour bounding boxes are used. However, this is not completely accurate, to know if a contour really covers an external defect, the algorithm must examine and adjust the width,

length, and width-length ratio of the region. To accomplish this, for each selected candidate rectangle, an extended region surrounding the curve is analyzed. The top and bottom boundaries of the enclosing rectangle are expanded each by a length of 10 cross sections (8 inches) along the log length. The reason an extended region surrounding the curve is analyzed is because often a curve only encloses the most-protruding portion of the defect, not the entire defect. Next, we determine the widest consecutive segment of each cross section within the region whose data points have radial distances greater than the contour level. Here a segment refers to a set of lines connecting the adjacent log-data points in the same cross section and enclosed in the contour curve. This step yields precise shape information about the potential surface-defect regions.

Using the shape information, some regions are identified as small, long strips of bark. All of these are rejected from further consideration if they are more than 25 $in^2$ and long and narrow (i.e., at least 75 percent of the segments in the contour have a ratio less than 0.8 between their widest consecutive segments and the total width of the region). Our expertise in external defect characteristics indicates that regions with such features are unlikely to be defective.

Because of limitations of our original data collection process, small regions that are too close to the top or bottom of a contour plot image are too ambiguous for analysis and thus are rejected as well. They either enclose partial defects, which the algorithm is incapable of detecting, or a small defect that cannot be accurately detected because of the current data resolution. This is likely an artifact of the original scanning process, and we do not identify defects near or outside the scanned area for testing purposes. For the remaining regions to be examined, we identify segments that are wide enough (width of the widest consecutive segment greater than 1/4 the width of the bounding rectangle). Thus, we can determine whether the top or bottom of an enclosed region is a narrow and long fragment, indicating bark, instead of being part of the actual defect. If such a fragment exists, the top or bottom boundary for the region is adjusted to remove the bark artifact. Then based on the adjusted width/length ratio and the adjusted size, the region might be rejected as being long and narrow, and thus not a defect.

Regions that are smaller than 50 $in^2$ and are located near larger candidates (less than 3.5 inches apart horizontally or vertically) are excluded. In such cases, the larger candidates more likely indicate the true defects, while the smaller ones are likely continuations of the same defect. Among candidates with good length (less than 7 inches), or length longer than 7 inches and width/length ratio greater than 0.2, those less than 50 $in^2$, and less than 3.5 inches apart from the selected larger ones, are excluded. Regions with areas less than 8 $in^2$, or less than 15 $in^2$ with a width/length ratio that is out of range (less than 0.5 or greater than 2), also are removed as they are too small to be recognized as a defect.

Next, candidates are then checked for missing data. If there are more than 20 points missing in a segment, that is, in the data cross section there is a gap wider than 1 inch, the segment is classified as corrupted. If there are more than 50 percent corrupted segments enclosed in the contour, the region is classified as severely missing data and is rejected.

A sawn top, or sawn branch stub, is a type of external defect where the tree limb was cut by loggers in the woods (**Fig. 4**). Therefore, the surface is often not completely leveled with respect to the log surface but tilted at an arbitrarily small angle. The sawn top is often not completely flat and sometimes ridges are present from the sawing process. Typically, part of the sawn branch stub will fall below the highest contour level, and this section of the defect needs to be recognized. Our algorithm is able to locate such regions using a "straight-line" segment technique described below and is capable of adjusting the boundaries to identify the entire flat-top region.

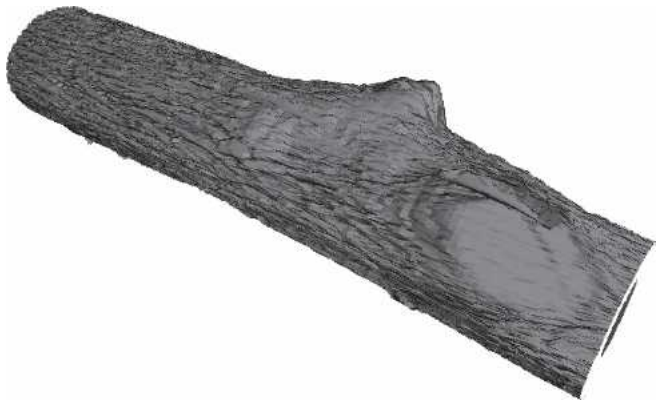

*Figure 4. — Graphically rendered view of log surface data showing a large branch stub.*

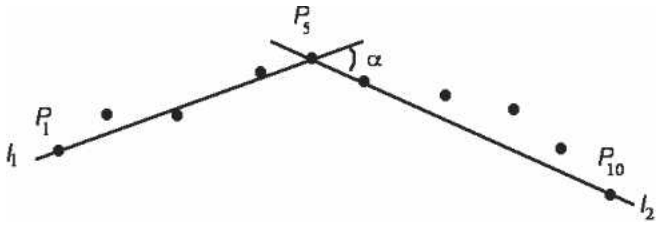


*Figure 5. — $l_1$ connects the first and fifth points, $P_1$ and $P_5$, respectively. $l_2$ connects the fifth and tenth points, $P_5$ and $P_{10}$. If the angle α between $l_1$ and $l_2$ is < 25 degrees, then the lines form a relatively straight line.*

For remaining regions with an area less than 25 $in^2$, the algorithm examines the angle changes between lines connecting log data points at an interval of five points along the cross sections (**Fig. 5**). If the changes are small enough (less than 25°), the corresponding segments are recorded as relatively straight. Then the range of "straight" segments is determined. If more than half of the segments contain straight parts, this region is identified as a sawn top, either sound (not rotten) or unsound (rotten). The boundary of the identified region is adjusted to surround all the regions containing these "straight" segments, so as to capture that portion of the sawn top that falls below the contour level.

Some regions may be falsely identified as a sawn top, because they contain severe missing data causing the algorithm to generate an incorrect result. Thus, they are rejected depending on the amount of missing data. Since the process of identifying sawn tops is often accompanied by adjustment of the defect region boundaries, which affects the geometric relationships among the detected regions, we again check for and remove regions that are completely nested or partially overlapped. To this point, those candidates that have survived are considered to be the most obvious and severe defect regions. Their rectangular borders are plotted on the contour image

and are labeled with their rank number in decreasing order of region areas.

So far, the algorithm has attempted to locate the most obvious defect types (Part 1 of the pseudo-code description). These defects comprise large bump-like knots, either old (healed broken stubs) or new (sawn at harvest). They may be large (20 inches diameter) or relatively small (4 inches diameter), protruding (at least 3 inches high) or with a more gentle rise, and sound or unsound. There is another group of severe defects with a medium rise (0.5 to 1 inch) and a medium diameter (3 to 5 inches). Because of these characteristics, the defects are not enclosed in the highest contour curves and thus are not identified by the procedure described so far. However, they have a distinctive pattern (surface rise and diameter). Thus, we provide an algorithm explicitly designed to identify these defects, which we refer to as less severe defects. In our sample of 14 logs, we observed eight such defects.

Next, the detection algorithm searches for the existence of upward slopes and downward slopes that meet the criteria composing a certain range of the gradient. A slope here refers to a group of adjacent data points whose radial distances increase or decrease along the log length in a general trend, similar to a slope in a mountain. During the process, a group of adjacent data points along the log length (z-axis) are examined. In this procedure, the type of defects we are looking for are not large or protruding—those defects should have been detected earlier. If the gradient falls within a certain range (high enough, but not so high as to represent a protruding defect that should have been detected in the first stage), it is tagged. Note that the predominant surface feature of a log is bark, which has an uneven texture. Therefore, the data points on a slope usually do not form a strict straight line. Our algorithm detects such slopes by judging their tendency, either going up or down, and an appropriate tolerance threshold is applied.

Based on the results from slope detection, those regions with: (1) width and length of 3 to 5 inches, (2) height of 0.5 to 1 inch, and (3) sufficient number of upward slopes and downward slopes are determined, and less severe defects identified. This kind of defect can also include rotten and non-rotten, sawn, or naturally formed defects. The detected less-severe defects are plotted in the same contour image with the severe defects previously identified. This completes the algorithm.

## Simulation results and discussion

Fourteen log data samples were chosen based on their data characteristics and analyzed using the defect detection system. The defect diagrams of all external defects present on log samples were collected manually by our sponsor, the USDA Forest Service in Princeton, West Virginia. Since logs are heavy (1,000 to 5,000 pounds) and vary in form (taper, sweep) and size (diameters at the two ends), accurately measuring the defect locations, sizes, and classifying defect types proved challenging. Consequently the diagrams are often erroneous, ambiguous, and inaccurate. Further, the diagrams often only record the width and length of a defect, but not its height or surface rise. External defects may not always be visible in the color images of a sample log, and the order of the color photographs were often incorrect. Among the 160 or so scanned log data samples, 45 of the logs were in poor quality and not usable.

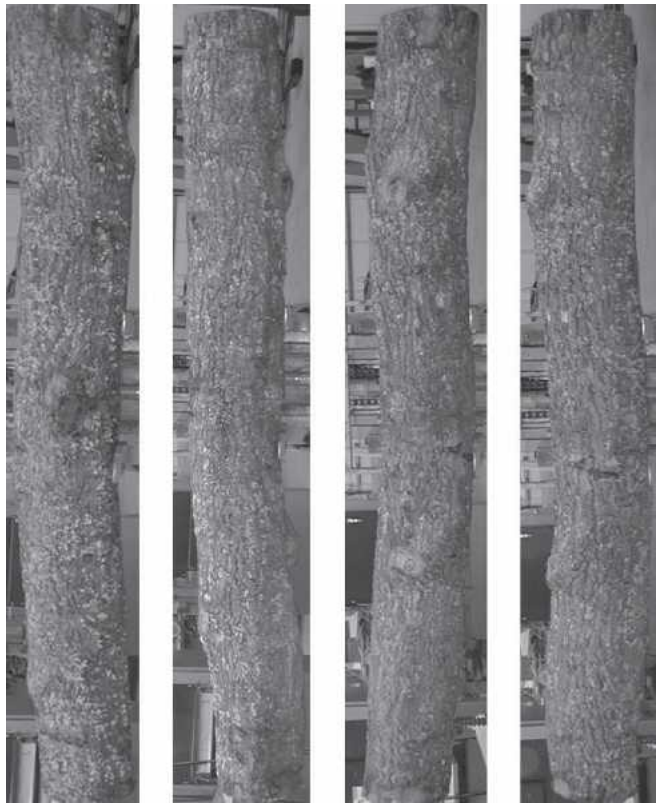

**Figure 6. — Digital images of the same log sample as in Figures 2 and 3, at 90° per side. These images are used, in part, to determine the correctness of the machine generated defective regions.*

The defect diagrams illustrate not only the defects visible in the radial-distance gray images, but also those undetectable as a result of the methods we adopted and/or the data resolution limits. We combined the information from the diagrams as well as from the color images (**Fig. 6**) and marked the observed defects in grayscale images that illustrate surface height variations (**Fig. 7**, right) and refer to them as ground truth. The coordinates of the marked rectangles are measured and recorded, which are then automatically overlaid on the contour plot (**Fig. 7**, left). In the contour plot, the observed defect regions are marked in solid crossed rectangles, while the automatically detected regions are marked in dashed crossed rectangles. The locations, widths, and lengths of automatically detected regions are reported by the program. To determine whether a marked region in the contour plot correctly indicates an external defect, we compare it with the ground truth. Among the 14 log samples, there are a total of 68 observed defects based on the grayscale image of radial distances, where 63 were correctly identified by the detection algorithm. Most non-identified defects are small (less than 5 inches in diameter) and/or relatively flat (less than 1 inch in surface rise). There are 10 non-defective regions falsely identified as defects. Nine of 10 false positive regions contain high-rise bark regions that are enclosed in the highest contour curves. Their widths and lengths range from 6 to more than 20 inches. The algorithm fails to remove them from the true defects using the criteria described in the previous section.

**Table 1** gives a breakdown for each log sample of observed defect counts, detected defect counts, falsely identified defect

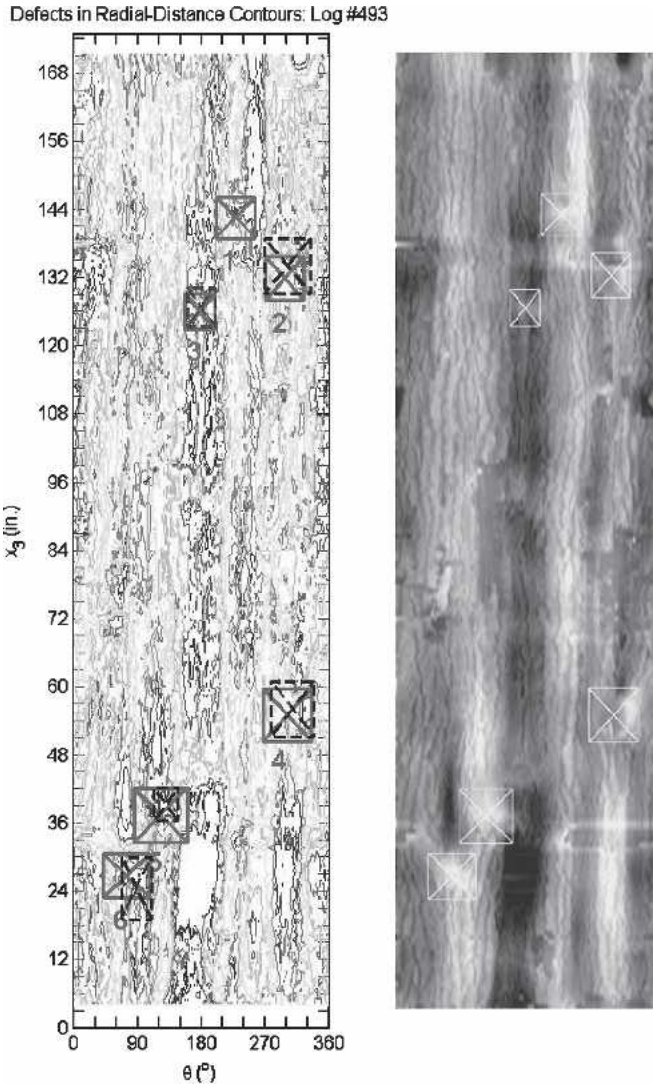


*Figure 7. — Left: a contour plot automatically generated by the defect detection programs. Dashed rectangles mark the possible defective regions, while solid rectangles are overlaid on observed defective regions. Right: The corresponding grayscale image with manually marked defect regions. Using the contour method, our algorithm finds five of six defects, where a match is defined when the center of an automated region falls inside the corresponding observed region, and vice versa.*

counts, and missed defect counts. **Table 2** gives a summary for each log sample of the surface area, detected defect area, falsely identified defect area, and missed defect area, all in square inches. The majority of the samples are red oak. There were no missed defects with the yellow-poplar samples. However, the number of samples for yellow-poplar is low. From **Tables 1** and **2**, we found that the average size of a correctly detected defect is 162 inch$^2$, but the average size of a missed defect is 51 inch$^2$. This tells us that the missed defects tend to be relatively smaller.

We used two methods to evaluate the performance of the detection algorithm. The first method counts the number of defects automatically detected out of the total number manually identified. In our experiments, there are a total of 68 severe defects of which 63 were correctly identified (**Table 1**).

*Table 1. — Counts for observed defects, detected defects, falsely identified defects, and missed defects for each log sample.*

| Log number | Species | Total | Correct | False positive | Missed |
|---|---|---|---|---|---|
| 444 | Red Oak | 4 | 4 | | |
| 448 | Red Oak | 9 | 8 | 1 | 1 |
| 450 | Red Oak | 4 | 3 | | 1 |
| 453 | Red Oak | 7 | 6 | | 1 |
| 468 | Red Oak | 3 | 3 | 1 | |
| 480 | Red Oak | 6 | 6 | | |
| 493 | Red Oak | 6 | 5 | | 1 |
| 501 | Red Oak | 3 | 3 | | |
| 508 | Red Oak | 5 | 5 | 1 | |
| 521 | Red Oak | 6 | 6 | 1 | |
| 537 | Red Oak | 5 | 4 | 2 | 1 |
| 441 | Yellow-Poplar | 2 | 2 | | |
| 485 | Yellow-Poplar | 6 | 6 | 2 | |
| 520 | Yellow-Poplar | 2 | 2 | 2 | |
| Total | | 68 | 63 | 10 | 5 |

*Table 2. — Log surface and surface areas for observed defects, detected defects, falsely identified defects, and missed defects for each log sample.*

| Log number | Species | Surface area | Observed defect area | Detected defect area | False positive area | Missed defect area |
|---|---|---|---|---|---|---|
| | | ---------- | ---------- | (in$^2$) | ---------- | ---------- |
| 444 | Red Oak | 5,797 | 456 | 456 | | |
| 448 | Red Oak | 7,284 | 1,196 | 1,105 | 30 | 91 |
| 450 | Red Oak | 7,278 | 570 | 553 | | 17 |
| 453 | Red Oak | 6,301 | 1,732 | 1,671 | | 61 |
| 468 | Red Oak | 5,453 | 959 | 959 | 122 | |
| 480 | Red Oak | 7,486 | 1,256 | 1,256 | | |
| 493 | Red Oak | 8,551 | 364 | 314 | | 50 |
| 501 | Red Oak | 3,916 | 445 | 445 | | |
| 508 | Red Oak | 4,031 | 573 | 573 | 243 | |
| 521 | Red Oak | 8,560 | 493 | 496 | 113 | |
| 537 | Red Oak | 6,414 | 390 | 356 | 178 | 34 |
| 441 | Yellow-Poplar | 4,645 | 297 | 297 | | |
| 485 | Yellow-Poplar | 9,352 | 1,385 | 1,385 | 309 | |
| 520 | Yellow-Poplar | 6,188 | 358 | 358 | 218 | |
| Total | | 91,256 | 10,474 | 10,224 | 1,213 | 253 |

There are 10 non-defective regions falsely identified as defects. Most non-identified defects are small (less than 5 inches in diameter) and/or relatively flat (less than 1 inch of surface rise). Nine of 10 falsely identified regions contain high-rise bark regions that are enclosed within the highest contour curves. Their widths and lengths range from 6 to over 20 inches. The algorithm fails to remove them from the true defects using the criteria described in the previous section.

The other way to evaluate the performance of the detection algorithm is to calculate the surface area of the detected defects against that of the ground truth. This is similar to the analysis used by Kline et al. (1998). We first estimate the total surface area of: log samples (LSA), which is 91,257.06 in$^2$; observed external defects (ODA), 10,474 in$^2$; automati-

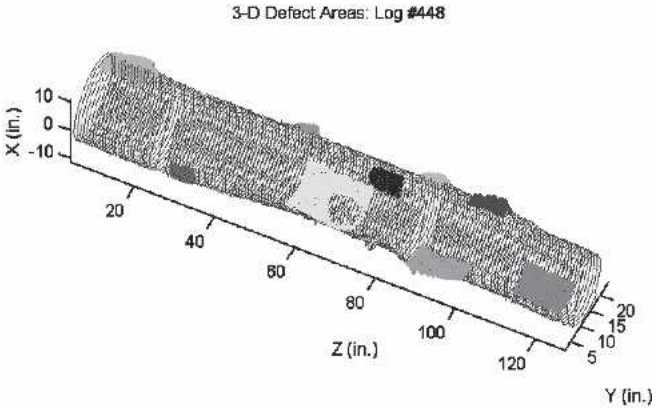


*Figure 8. — A 3-D rendering of the log data with automatically detected defects marked by patches. Such an image might be used by sawyers to maximize the value of wood products.*

cally identified defects that match the observations (MDA), 10,224 $in^2$; automatically identified defects that do not match the observations (FPA), a false positive, with a total of 1,213 $in^2$. We also define ADA as the total area of all defects determined by the detection algorithm, which equals 11,435.21 $in^2$, and FNA, the total area of all observed defects that are NOT identified by the detection algorithm—false negative, 253.39 $in^2$.

When the centerpoint of a detected region falls inside the bounding box of an observed defect, and vice versa, we declare it a match and use the defect area given by the ground truth in calculation. Now we obtain the detection statistics: the percentage of observed clear region is 88.52 percent ((LSA-ODA)/LSA × 100%). The percentage of automated clear region is 87.47 percent, given by (LSA-ADA)/LSA × 100 percent. That the latter is smaller than the former implies that the algorithm identified more defective surface area than the actual observed area. The percentage of false positive or the falsely identified defect regions from clear surface is 1.50 percent (FPA/(LSA-ODA) × 100%). The percentage of false negative, indicating how much the algorithm missed the defective regions, amounts to 2.42 percent (FNA/ODA × 100%). Finally, 97.58 percent is the detection rate for our defect detection algorithm with respect to observations, given by MDA/ODA × 100 percent. Since the total of FNA and MDA is equivalent to ODA, the false negative rate and the detection rate add up to 1.

## Future work

The Matlab code that detects defects will be converted to Java and integrated with the scanning and sawing equipment. The detection results will be displayed in graphical formats to assist sawyers who can rotate, zoom, and move the virtual log marked with defects (**Fig. 8**). We implemented the method to compute the false-detection rate as discussed in the previous section, which demonstrated a reasonably good algorithm (87.47% automated clear region vs. 88.52 % observed clear region, and a 97.58% detection rate). Many defects were not identified mainly because they do not have a significant height change. Thus, our contour approach is not effective for these defects. Among them there is a group of defects that are severe, e.g., heavy distortions and flat knots. These defects often have a distinctive ring-like bark pattern. Edge detection, a computer vision technique, may help in identifying such defects. This will be investigated in the second phase of this research.

## Literature cited

Carpenter, R.D., D.L. Sonderman, and E.D. Rast. 1989. Defects in Hardwood Timber. Ag. Handbook 678. U.S. Dept. of Agriculture, Washington, D.C. 88 pp.

Chang, S.J. 1992. External and Internal Defect Detection to Optimize Cutting of Hardwood Logs and Lumber. Transferring Technologies for Industry No. 3. USDA and National Agri. Library, Beltsville, Maryland.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel. 1986. Robust Statistics: The Approach Based on Influence Functions. John Wiley. New York.

Kline, D.E., A. Widoyoko, J.K. Wiedenbeck, and P.A. Araman. 1998. Performance of color camera machine vision in automated furniture rough mill systems. Forest Prod. J. 48(3):38-45.

Li, P., A.L. Abbott, and L.S. Daniel. 1996. Automated analysis of CT images for the inspection of hardwood log. *In*: Proc. of the Inter. Conf. on Neural Networks. Washington, D.C., USA. Vol. 3:1744-1749.

Thomas, E., L. Thomas, C.A. Shaffer, and L. Mili. 2006. Using external high-resolution log scanning to determine internal defect characteristics. Submitted to the 15th Central Hardwood Forest Conf. USDA Forest Serv., SE Res. Sta., Asheville, North Carolina. (In process.)

_____, _____, L. Mili, R. Ehrich, A.L. Abbott, and C.A. Shaffer. 2003. Primary detection of hardwood log defects using laser surface scanning. *In*: ISandT/SPIE Electronic Imaging 2003, 20-24 January 2003, Santa Clara, California. Vol. 5011:39-49.

Thomas, L. and L. Mili. 2006. A Robust GM-Estimator for the automated detection of external defects on barked hardwood logs and stems. Accepted by IEEE Transactions on Signal Processing. (In process.)

Thomas, L., L. Mili, C.A. Shaffer, and E. Thomas. 2004. Defect Detection on Hardwood Logs Using High Resolution Three-Dimensional Laser Scan Data, IEEE ICIP 2004, Singapore, October 24-27. pp. 243-246.

Zhu, D., R. Conners, F. Lamb, and P. Araman. 1991. A computer vision system for locating and identifying internal log defects using CT imagery. *In*: Proc. of the 4th Inter. Conf. on Scanning Tech. in the Wood Industry. Miller Freeman Publishing, Inc., San Francisco, California. pp. 1-13.