DIFFERENTIATING DEFECTS IN RED OAK LUMBER BY DISCRIMINANT ANALYSIS USING COLOR, SHAPE, AND DENSITY

Brian H. Bond

Assistant Professor
Department of Forestry, Wildlife and Fisheries
University of Tennessee
P.O. Box 1071
Knoxville, TN 37909

D. Earl Kline†

Professor
Department of Wood Science and Forest Products
Virginia Polytechnic and State University
Blacksburg, VA 24061-0503

and

Philip A. Araman

USDA Forest Service, Southern Research Station
Brooks Forest Products Center
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061-0503

ABSTRACT

Defect color, shape, and density measures aid in the differentiation of knots, bark pockets, stain/mineral streak, and clearwood in red oak, (*Quercus rubra*). Various color, shape, and density measures were extracted for defects present in color and X-ray images captured using a color line scan camera and an X-ray line scan detector. Analysis of variance was used to determine which color, shape, and density measures differed between defects. Discriminant classifiers were used to test which defect measures best discriminated between different defects in lumber.

The ANOVA method of model measure selection was unable to provide a direct method of selecting the optimum combination of measures; however, it did provide insight as to which measure should be selected in cases of confusion between defects. No single sensor measure provided overall classification accuracy greater than 70%, indicating the need for multisensor and multimeasure information for defect classification. When used alone, color measures resulted in the highest overall defect classification accuracy (between 69 and 70%). Shape and density measures resulted in the lowest overall classification accuracy (between 32 and 53%); however, when used in combination with other measures, they contributed to a 5–10% increase in defect classification accuracy. It was determined that defect classification required multisensor information to obtain the highest accuracy. For classifying defects in red oak, sensor measures should include two color mean values and two standard deviation values, a shape measure, and a X-ray standard deviation value.

Keywords: Lumber scanning, defect detection, discriminant analysis, machine vision.

INTRODUCTION

With a shortage of qualified labor, a desire to use lower grade raw materials, and an increased yield benefit of using automation, there is a growing interest in developing and using automated defect detection systems. These systems are of particular interest to the hardwood manufacturing industry where defects must be removed from visible faces as in flooring, furniture, cabinets, and millwork. Automated defect detection allows for more

† Member of SWST.

Wood and Fiber Science, 34(4), 2002, pp. 516–528

precise production control methods and production concepts such as multiple batch processing and the inclusion of character-marked material.

It has been proposed that new automation technologies will require the use of multiple sensor scanning systems (Kline et al. 1993). Although different types of sensors have been studied, little work has been published on the combination of different sensors in one system (Conners et al. 1992; Åstrand 1992; Hagman and Grundberg 1993). By integrating multiple sensor information, the accuracy of an automatic lumber inspection system can be substantially improved. While many defect detection systems have been able to differentiate between certain defect types, little information is known about what combination of sensors and what sensor information achieve the most accurate differentiation between defect types.

Defects in lumber are commonly observed by visual examination; however, other indications, such as density, can also be used to differentiate between clearwood and defects. Many different sensing methods have been applied to inspection of wood including optical, ultrasonic, microwave, nuclear magnetic resonance, and X-ray sensing (Bond 1998). Most automated defect detection systems require either visual (color, light intensity, and/or shape), or density differences to differentiate between defect types.

Conners et al. (1985), Silven and Kauppinen (1996), and Brunner et al. (1990) demonstrated the utility of color information for defect detection in wood. Color is an "attribute of visual perception that can be described by color names such as white, gray, black, yellow, orange, brown, red, green, blue, purple, etc., or by combinations of such names" (Grum and Bartleson 1980). The color of a material is determined by the spectral makeup of light reflected from its surface. Due to the limited understanding of the human visual system, many methods of describing or modeling of color exist. The most common and simple is the RGB color model, which uses three primary colors (red, green, and blue) to describe a color within a color range (Weeks 1996). The RGB model describes a color image as a set of three independent gray-scale images each having 256 gray levels. Other color models exist and are described by Weeks (1996). The use of various color models to describe the color of wood has been studied by Brunner et al. (1990); however, the RGB color space is the most common.

Little detailed information exists in literature describing the color differences of individual wood features. Most information on the subject is based on the selection of color spaces for the classification of all defects rather than how each feature can be described by a particular color space. Performance of a color space is based on the classification accuracy of all feature types. By better understanding how each feature is represented by color measures, improved classification performance could be achieved.

The need for density information to improve defect detection accuracy has been documented by several researchers (Conners et al. 1990, 1992; Araman et al. 1992; Portala and Ciccotelli 1992). Differences in density have long been used in the nondestructive evaluation of objects. One common method of evaluating density differences within a material is through the use of X-ray attenuation. The attenuation of radiation through an object to detect features based on density differences is known as radiography. Portala and Ciccotelli (1992) demonstrated the ability of X-ray information to be used in identifying features in wood that contain density differences.

Several investigators have demonstrated the usefulness of shape measures in the detection of wood defects (Forrer et al. 1988; Pölzleitner and Schwingshakl 1990; Cho 1991; Kauppinen and Silven 1995; Lampinen and Smolander 1996). The most common shape measures for defect recognition include area, eccentricity, Euler number, compactness, and the slope density function (Ballard and Brown 1982). Lampinen and Smolander (1996) stated that while shape information is of great importance

TABLE 1. *Definition of prevalent wood defects included in this study.*

Defect	Definition
Bark pockets:	Inclusion of bark within the wood where a knot or wane are not present.
Knots:	Part of the limbs that are embedded in the main stem.
Stain (1) and mineral streak	(1) Discoloration caused by fungi or bacteria. Initial evidences of decay. (2) An olive to greenish-black or brown discoloration of undetermined cause in hardwoods.
Clearwood:	Normal earlywood and latewood area that contains no other defects

for defect classification of wood features, it is insufficient when used alone.

While the scanning methods discussed are useful in systems that detect wood features, no single method has proven adequate for all defect types. By optimally combining several measures from multisensor scanning systems, a more reliable defect detection system can be developed, However, it is not understood how well these different measures can describe individual lumber defects. By gaining knowledge of what feature measures are important for classification, more appropriate sensing methods can be selected and better defect detection methods can be developed.

The objective of this research was to determine the differences between color, shape, and density measures between various wood defects (features) and to determine the best combination of measures for classification. Defects studied were knots, stain, bark pockets, and clearwood in red oak lumber. Analysis of variance and discriminant analysis techniques were used to study how different combinations of measures performed in differentiating between the various defects.

METHODS

Materials and measurements

Images of clearwood, knots, bark pockets, and stain were attained from red oak, (*Quercus rubra*) lumber samples. Images were collected using a color line scan camera and an X-ray line scan detector mounted on equipment designed for multisensor scanning of lumber (Conners et al. 1997; Bond 1998). Red oak lumber was chosen because of its commercial importance to the furniture and cabinet industry. Eight-foot samples of No. 2 Common lumber (National Hardwood Lumber Association, NHLA 1994) were obtained in the kiln-dry condition (8% MC) from various sources within the Appalachian mountain region. The widths of the samples were less than 10 in. Moisture content of 8% was chosen because it is commonly used in the manufacturing of wood products in the United States and allowed for stable moisture conditions in the laboratory. All lumber was planed within 7/8 in. to 15/16 in. to create a clean surface free of soil, grease, surface roughness, and other marks that would alter the natural variability of the features studied.

The defects investigated in this study are listed in Table 1. Defects were selected based on their frequency and the area they occupy in NHLA graded No. 1 and No. 2 Common red oak (Buehlmann et al. 1997; Wiedenbeck and Buehlmann 1995). Only intergrown (sound) knot specimens were included in this study. Twenty samples of clearwood, knots, and bark pockets, and nineteen samples of stain/mineral were selected from the lumber.

The lumber was scanned using a multiple sensor detection system assembled at Virginia Tech (Fig. 1), which generates both X-ray and color images with a cross-board resolution of 1.2 pixels/mm (30 pixels/in.) and a down-board resolution of 0.63 pixels/mm (16 pixels/in.) (Conners et al. 1997). The lumber was scanned at a rate of one foot per second. The X-ray and color systems were configured to have identical spatial resolution and alignment.

Color images were collected using a Pulnix TL-2600 RGB line scan camera with a line

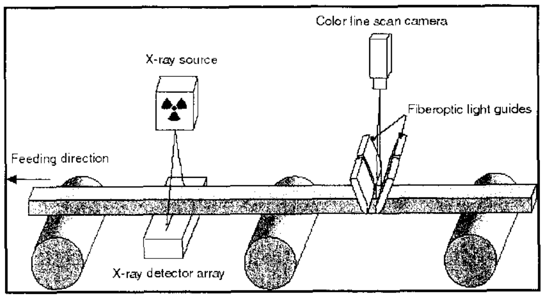

FIG. 1. Lumber scanning system.

resolution of 864 pixels. The camera was mounted perpendicular to the wood surface, and four linear Fostec fiber-optic light lines were used to illuminate the surface of the lumber. These light lines were arranged pairwise in a small angle from the optical axis of the camera illuminating a line across the boards perpendicular to the feeding direction (see Fig. 1). The Pulnix camera has three sensor arrays filtered with red, green, and blue interference filters, respectively. This camera configuration gives a slight spatial offset of the three color images, but no undesired effects were observed for this experiment. The camera was fitted with two color-balancing filters (Schott filter numbers FG-6 and BG-34), and the three color channels were individually shade-corrected with a linear function.

The X-ray scanning system provides an averaged image of the wood density throughout the lumber thickness. This average wood density is useful for detection of defects in wood since most defects are of higher or lower density than normal or clear wood. The X-ray system employed an EG&G Astrophysics X-ray source with the radiation energy set to 100 keV and 0.6 mA. The X-ray sensor was a 256-pixel line array manufactured by FISCAN. The images were shade-corrected using a linear function, and the contrast was optimized for 4/4 red oak by calibrating the minimum level (highest absorption) with a target of 45-mm-thick polyethylene.

Wood feature images were imported into Image-Pro Plus image analysis software (Media Cybernetics 1995). The area contained within a defect was manually isolated. The exact location and size of a defect were measured to verify defect locations in the images. Measurements of color (mean and standard deviation for each of the red, green, and blue image gray-scale values), shape (aspect and roundness), and X-ray attenuation (mean and standard deviation for the X-ray image gray-scale values) were quantitatively measured for the selected defect region. The measurements collected from each defect region, their notation, and definitions are presented in Table 2. Aspect and roundness measurements were calculated from measurements taken from the color image. Aspect is the ratio of maximum cord A to maximum cord B perpendicular to A as shown in Fig. 2. Roundness of an object is defined as

$$\text{Roundness} = \frac{\rho^2}{(4 \times \rho \times a)} \quad (1)$$

TABLE 2. *Definition of properties used to characterize wood defects.*

Property	Notation	Feature measured from each defect region	Image used to measure
Color	R_m	Red mean	Color
	G_m	Green mean	
	B_m	Blue mean	
	R_s	Red standard deviation	
	G_s	Green standard deviation	
	B_s	Blue standard deviation	
Shape	ASP	Aspect	Color
	RND	Roundness	
X-ray attenuation	X_m	Attenuation mean	X-ray
	X_s	Attenuation standard deviation	

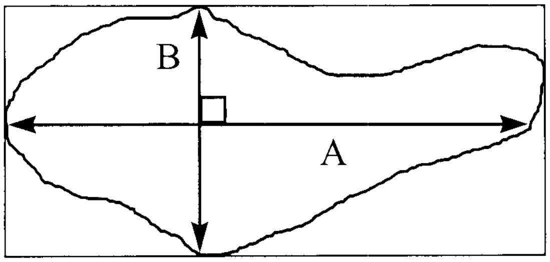

FIG. 2. Eccentricity measure defined A/B.

where ρ is the perimeter (cm), and a is the area (cm^2).

While the X-ray image is actually a grayscale representation of the attenuation of X-rays through the wood, it is also a representation of the relative density through the thickness of the material. Hence, this measure will be referred to as the density measure in further discussion. To collect the density measures that spatially correspond to the color measures, the defect area was manually segmented from clearwood, measured, and then superimposed on both the color and X-ray images to insure that the same region in the lumber was selected for measurement.

Statistical procedures

The relationship of the color, shape, and density measures between defects was determined using analysis of variance techniques (Bond 1998). The assumption of equal variance for the tested groups was verified before each ANOVA. Tukey's W procedure was used to determine which groups were significantly different. Tukey's W procedure was chosen since it is conservative in determining group differences and is able to account for unequal sample sizes (Ott 1988). SAS (SAS Institute 1996) statistical analysis software was used for all statistical analysis.

To study the effectiveness of the discovered differences, color, shape, and density measures were used as classification variables in a discriminant analysis. All variables (measures) were tested individually to determine their ability to separate defects. Then, only those measures that were shown to be significantly different for multiple defect types were used as classification variables. Finally, variables were systematically added and deleted to determine their effect on the classification accuracy.

Discriminant analysis procedures

Discriminant functions are often used in pattern recognition and classification for the separation of data into groups (Duda and Hart 1973). Discriminant analysis is a statistical technique that is used to classify objects into groups based on *a priori* information. The use of *a priori* information is a limiting factor of this work as drastically different color or density in samples could change the results.

The most common methods of discriminant analysis include Fisher's Discriminant function and linear logistic discrimination (Everitt and Dunn 1992). Other methods are available and are reviewed in Hand (1981). The many different methods are a result of the variety of distribution assumptions made about the variables describing the feature to be classified; however, Fisher's linear discriminant function has been shown to be relatively robust for those situations where there is a departure from normality (Everitt and Dunn 1992). Fisher's approach to constructing a classification rule is based on specifying a theoretical probability distribution model (normal) and assuming that the data fit the model, then estimating the parameters using the data, and finally constructing a rule using these estimates (Everitt and Dunn 1992).

Discriminant analysis techniques used in this study utilized a group of classifiers referred to as quadratic classification functions. In general, each observation is placed in the class from which it has the smallest generalized squared distance. For example, an observation f is assigned to class c if

$$Q_{fc} > Q_{fc'} \qquad (2)$$

where Q_{fc} is the quadratic classification function score for observation f in class c, and $Q_{fc'}$ is the quadratic classification function score for observation f in any other class c'. The quadratic classification function score is calculated as

$$Q_{fc} = \ln q_c - \tfrac{1}{2}\ln|S_c| - \tfrac{1}{2}D^2_{fc} \quad (3)$$

where q_c is the estimated prior probability for class c, S_c is the covariance matrix for class c, and D^2_{fc} is defined as sample distance between observation f and class c.

When the within-group covariance is used, the distance from observation f to class c is

$$D^2_{fc} = (X_f - \bar{X}_c)'S^{-1}_c(X_f - \bar{X}_c) \quad (4)$$

where D^2_{fc} is the sample distance between observation f and class c, S_c is the covariance matrix with class c, X_f is a p-dimensional vector containing the quantitative or measured features of an observation, X_c is the p-dimensional vector containing the feature parameters that are known to represent class c.

The misclassification rate of the discriminant function can be determined several ways as described in Hand (1986). The most common method is the "leaving one out method" where the discriminant function parameters are derived on the basis of (n − 1) subjects and is used to classify the individual or observation not included (Everitt and Dunn 1992). The whole process is then repeated for each individual to be classified.

The McNemar chi-squared statistic is used to compare the overall classification accuracy of two different rules (Huberty 1994). The McNemar chi-squared statistic compares the proportion of correctly classified defects of one rule to the portion of correctly classified defects of another.

For this study there were 4 defect classes and sample sizes from 19–20 available for estimating parameters for most defect classes. The number of feature measures used to represent a class ranged from 1 to 10. Discriminant classifiers using 10 measures were close to overfitting the data. Since all of the classifiers tested used no more than 7 classification variables, the level of samples fell within an acceptable level for adequate discriminant analysis for the highest performing classifiers. The sample size of 19–20 for each defect category is relatively small for quadratic discriminant classification and is a limiting factor of the study.

TABLE 3. *Accuracies for single sensor vs. multi sensor classifiers.*

	Classification accuracy (percent classified correct)					
Variable	Bark	Clearwood	Knots	Stain/ mineral	Overall	a
R_m	95	100	65	21.1	70.9	A[a]
G_m	95	100	60	21.1	69.6	A
B_m	95	100	60	21.1	69.6	A
X_m	20	65	45	0.0	32.9	B
ASP	45	10	90	47.4	48.1	C
RND	65	0	95	52.6	53.2	C
R_s	40	50	60	0.0	37.9	DB
G_s	0	75	85	15.8	43.3	DC
B_s	20	85	80	5.3	48.1	C
X_s	20	80	50	31.6	45.6	C

[a] Where comparisons of two classification accuracies share the same capital letter, no significant difference exists at the α = 0.01 level.

The methods employed in this study involve the use of shape measures in both ANOVA and discriminant analysis. When a shape measure was tested in the analysis, it was included to represent all defect classes. While shape measures are not applicable for clearwood, it is important to classify suspect defect regions as clearwood when they are, in fact, normal or defect-free wood. To eliminate bias toward any arbitrary shape measures used in the clearwood samples, an average of all the other defect type shape measures were used for clearwood in all analysis.

RESULTS

Differentiation of defects

Discriminant classifiers were first developed and tested for each measure (Table 2) separately to evaluate their effectiveness as classification variables. The classification results from individual measures are presented for each defect in Table 3. The greatest classification accuracy for any one measure does not exceed 70% for any defect. The inability of single measure to achieve high classification accuracy (e.g., greater than 90%) demonstrates the need for multisensor and multimeasure defect classification methods.

TABLE 4. *Measures that have significant differences between defects in red oak.*

	Knot		Bark pocket		Stain/mineral	
Defect type	Measure	*P*-value	Measure	*P*-value	Measure	*P*-value
Bark pocket	R_m	0.010				
	G_m	0.010				
	B_m	0.001				
	RND	0.030				
	R_s	0.006				
Stain/mineral	ASP	0.001	R_m	0.001		
	RND	0.001	G_m	0.001		
	G_s	0.019	B_m	0.001		
	B_s	0.003	ASP	0.001		
	X_s	0.001	RND	0.001		
	R_s	0.068	X_s	0.003		
Clearwood	R_m	0.001	R_m	0.001	R_m	0.001
	G_m	0.001	G_m	0.001	G_m	0.001
	B_m	0.001	B_m	0.001	B_m	0.001
	X_m	0.009	X_s	0.001	ASP	0.001
	G_s	0.012			RND	0.001
	B_s	0.001			B_s	0.001
	X_s	0.001				

Table 3 is also useful for identifying where certain measures may provide information to improve classification accuracy. For example, classifiers based on mean color parameters give high accuracy (95% or higher) for both clearwood and bark, indicating that color measures are useful in differentiating these defect types.

To gain knowledge of how color, shape, and density measures differ between defects in red oak, the feature measures described in Table 2 were compared using Tukey's W analysis of variance. By characterizing how defects differ by sensor measures, it can be determined how these defects are unique and how to best differentiate between them. When a measure was found to be significantly different between two or more defects, it was suggested that the measure would be suitable for use as a classification variable for differentiating between the defect types. The use of ANOVA as a method to select classification variables was suggested by Huberty (1994).

The significant p-values (i.e., values less than 0.01) resulting from Tukey's multiple comparisons between defect types for red oak are listed in Table 4. All values listed as 0.001 are equal to, or less than this value. The table is in matrix form, where the defects are listed in the top row and left column, and the p-values representing the significant difference levels between the specific defect types are presented in the appropriate row and column. While the ANOVA results in Table 4 give an indication as to what measures are important to differentiate defects, they do not quantify the marginal contribution of individual measures to classification accuracy. The effectiveness of the ANOVA methods to select discriminant variables will be discussed in the following sections referring to the results of single measure (Table 3) and multimeasure (Table 5) classification accuracies.

Color measures

Table 4 shows the occurrence of high significance levels in color means, R_m, G_m, and B_m, for all defect types except between knot and stain/mineral. This occurrence indicates that these color measures are useful for differentiating between all defect types except between knots and stain/mineral. Table 3 shows that each of the color means, when used

TABLE 5. *Classification results for each variable for red oak defects.*

Variables included in classifier	Bark	Clearwood	Knot	Stain/mineral	Overall	[a]
Color measures						
R_m B_m	80.0	100.0	70.0	47.3	74.6	A
R_m G_m	95.0	100.0	70.0	21.1	71.5	A
G_m B_m	90.0	100.0	65.0	26.3	70.3	A
R_m G_m B_m	80.0	100.0	70.0	47.3	74.6	A
R_s G_s B_s	85.0	90.0	75.0	47.3	74.6	A
R_m R_s G_m G_s B_m B_s	100.0	95.0	90.0	84.1	92.4	B
R_m R_s B_m B_s	100.0	100.0	85.0	79.0	91.1	B
Shape measures						
ASP RND	65.0	40.0	95.0	52.6	63.3	C
ASP	45.0	10.0	90.0	47.4	48.1	D
RND	65.0	0.0	95.0	52.6	53.2	E
X-ray measures						
X_m X_s	25.0	75.0	45.0	31.6	44.2	D
X_s	20.0	80.0	50.0	31.6	45.6	D
X_m	20.0	65.0	45.0	0	32.8	F
Color and shape measures						
R_m R_s B_m B_s RND	90.0	100.0	95.0	94.7	94.9	B
Color and X-ray measures						
R_m R_s G_m G_s B_m B_s X_s	100.0	95.0	90.0	84.2	92.4	B
R_m R_s B_m B_s X_s	90.0	100.0	90.0	94.7	93.7	B
R_m R_s B_m B_s X_m X_s	85.0	100.0	90.0	94.7	92.4	B
Shape and X-ray measures						
ASP RND X_m X_s	75.0	60.0	95.0	52.6	70.9	AC
Optimum measure combination						
R_m R_s B_m B_s RND X_s	100.0	100.0	100.0	95.0	98.8	F

[a] Where comparisons of two classification accuracies share the same capital letter, no significant difference exists at the $\alpha = 0.01$ level.

individually as a classifier, gives similar accuracy results indicating that any one of the color means has equal importance towards classification. Since color means are not significantly different between knots and stain/mineral, they will not be effective in differentiating stain/mineral from knots. This difficulty is verified in Tables 3 and 5 where accuracy is lowest for stain/mineral and knot detection. No evidence was found to support that the use of more than two color means increased classification accuracy.

Table 4 shows that there were no prominent trends in the significant differences between defect types for color measure standard deviations. No one measure proved to be significantly different for all defect types. Table 3 verifies the lack of significance between defect types for color standard deviation measures since when used individually, all resulted in low classification accuracies (e.g., 47.6% or less). However, when at least two color standard deviations are combined in a classifier (Table 3 and Table 5), the overall accuracy increases significantly, indicating that better defect detection can be obtained by combining two or more color standard deviation measures as classification variables.

Table 4 shows that knots and stain/mineral were not significantly different for color means, but were significantly different for color standard deviations. Hence, by combining both color means and color standard deviations into a classifier, the accuracy of differentiating knots from stain/mineral should improve. Table 5 illustrates this improvement by

TABLE 6. *Accuracy of the top five classifiers and the classifier constructed of measures selected using ANOVA methods.*

Variables included in classifier	Bark	Clearwood	Knot	Stain/mineral	Overall	[a]
Top five classifiers						
R_m R_s G_m B_s ASP X_m X_s	95.0	100.0	90.0	100.0	96.3	A
R_m R_s B_m B_s ASP X_s	95.0	100.0	95.0	100.0	97.5	AB
R_m R_s B_m B_s ASP RND X_s	95.0	100.0	95.0	100.0	97.5	AB
R_m R_s B_m B_s RND X_m X_s	90.0	100.0	95.0	100.0	97.5	AB
R_m R_s B_m B_s ASP X_m X_s	95.0	100.0	95.0	100.0	97.5	AB
R_m R_s B_m B_s RND X_s	100.0	100.0	100.0	95.0	98.8	B

[a] Where comparisons of two classification accuracies share the same capital letter, no significant difference exists at the α = 0.01 level.

showing an overall classification accuracy increase (92.4% vs. 74.6%) along with a significant increase in the detection accuracy of knots (90.0% vs. 70%) and stain/mineral (84.1% vs. 47.3%). This result indicates that the proper combination of different measures from the same sensor can improve defect classification accuracies.

Density measures

High and low density defect types (e.g., knots and clearwood) were significantly different for X_m and X_s. However, defects with similar densities (e.g., stain/mineral and clearwood) were not significantly different for density measures. While these results indicate that X-ray information alone is not capable of unambiguously differentiating the four defect types in red oak (see Table 5), when used in conjunction with color information, they should improve the ability of differentiating between knots and stain/mineral since there is a significant difference in density between these two features. X_s was found to have a higher occurrence of significant differences between defect types than X_m indicating that it would be a more useful measure to include in defect classification. The results in Table 5 confirm this finding since X_s provides significantly higher classification accuracy than X_m.

Shape measures

ANOVA results show that RND is moderately better suited as a classification variable and Table 3 shows that RND provides higher overall classification accuracy than ASP. However, when combined with other variables, shape measures appear to be interchangeable without significantly affecting the overall classification accuracy (Table 6 and Table 7). The addition of a shape measure to a multimeasure classifier improves the overall accuracy (Table 5). Earlier it was noted that RGB color measures could be used to differentiate bark pockets and clearwood but not knots and stain/mineral streak. Since ASP and RND are significantly different for these two defect types, including a shape measure in a multimeasure

TABLE 7. *Contribution of shape measures to classification accuracy.*

Variables included in classifier	Bark	Clearwood	Knot	Stain/mineral	Overall	[a]
Contribution of shape measures:						
R_m G_m B_m B_s ASP RND X_s	80.0	100.0	90.0	94.7	91.1	A
R_m G_m B_m B_s RND X_s	90.0	100.0	90.0	100.0	94.9	A
R_m G_m B_m B_s ASP X_s	80.0	100.0	100.0	100.0	94.9	A
R_m G_m B_m B_s X_s	60.0	100.0	80.0	89.5	82.2	B
R_m R_s B_m B_s ASP RND X_s	95.0	100.0	95.0	100.0	97.5	C
R_m R_s B_m B_s RND X_s	100.0	100.0	95.0	100.0	98.8	C
R_m R_s B_m B_s ASP X_s	90.0	100.0	100.0	100.0	97.5	C
R_m R_s B_m B_s X_s	90.0	100.0	90.0	94.7	93.7	A

[a] Where comparisons of two classification accuracies share the same capital letter, no significant difference exists at the α = 0.01 level.

classifier will likely improve the ability of the classifier to differentiate between these defect types. Table 5 shows that adding RND with R_m, R_s, B_m, and B_s as classifiers, the accuracy of stain/mineral detection can be increased by almost 16% (79.0% vs. 94.7%). These results demonstrate that in red oak, defect differentiation will require a multimeasure or multisensor classifier to achieve the highest possible classification accuracy. The requirement of feature measures other than color in defect classification has been suggested by Lampinen and Smolander (1996) and Conners et al. (1990), who suggest the use of X-ray imaging for a density measure.

Combined measures

The classification results of discriminant functions with single variables demonstrate the need for multifeature measures. Further evidence to support this finding can be found by comparing combinations of individual sensor measures to those with multiple sensor measures. The highest obtainable classification accuracy of any combination of color measures from the same sensor type is no greater than 92.4% (see Table 5). The highest obtainable classification accuracy for X-ray combinations and shape combinations is 45.6% and 63.3%, respectively (see Table 5). The best combination of any three sensor measures increases the classification accuracy to 98.8%, which is 6.5% higher than color measures combined, 53.4% higher than obtainable with X-ray measures, and 35.6% higher than obtainable with shape measures alone (see Table 5). The increased classification accuracy obtained by combining measures helps to demonstrate that a combination of different measures provides greater accuracy. These results also indicate the importance of color in feature classification methods and that while color is able to give high classification results, the addition of other measures helps to reduce confusion in differentiating between difficult feature types.

The statistical relationships for defects found in the ANOVA analysis were used to guide the selection of measures for classification. Over 150 measure combinations were tested in discriminant functions with classification accuracy's ranging from 38 to 98.8%. The feature measures and accuracies for the top five classifiers are listed in Table 6. While the ANOVA method of measure selection provides some useful insight as to how individual measures differ between defects, it does not give any direct indication as to the optimum combination of measures for classification. Nevertheless, the ANOVA results do give an indication of the measures that should be included in cases where classification accuracy is limited by confusion between two defect types. For example, the ANOVA results (Table 4) show that R_s is significantly different between knots and stain/mineral, and the success of R_s in reducing classification confusion between these defect types is noted by the inclusion of R_s in the top five classifiers (Table 6).

When testing the measure combinations in classifiers, several important observations were made. A combination of all measures, while providing relatively high classification accuracy, tends to over-fit the data, indicating that it is important to consider which measures are complementary and which are over-fitting. For example, it was determined that only two color means are required to achieve the highest classification accuracy and that the inclusion of R_m in the classifier was also required to attain the highest possible accuracy (see Table 6). Reducing the number of measures or variables in defect differentiation can be very important to the speed of the classifier in industrial settings and also reduce the cost of a defect detection system.

Both shape and density measures were found to further increase performance of the discriminant classifiers. By comparing the classification accuracy of discriminant functions without shape measures to those with shape measures, it was discovered that including a shape measure could increase the classification accuracy for knots by 10%, stain/

mineral by 15%, and overall accuracy by 3.9%. The difference in classification accuracy for individual defects varied no more than 5%, based on which shape measure was used and, when used in a multimeasure classifier, there was no statistically significant difference in classification accuracy. While density alone does not provide high classification accuracy for all defect types, when included in a multimeasure model, it was found to reduce the confusion of knots and stain/mineral streak in the final model by 5% (Table 5).

While the high overall classification accuracy of 98.8% can be achieved, it was observed that some defect types are more difficult to classify than others. Most misclassifications involved stain/mineral streak. For example, classifiers had the most difficulty with bark pockets and knots, both of which were misclassified as stain/mineral streak. Stain has also proved a difficult defect to classify by other investigators, due to the large variability within the defect and its color similarity with other defects (Adel et al. 1993).

DISCUSSION AND CONCLUSIONS

It must be noted that the classification accuracy obtained in these results is fairly high. This is in part due to the classification being based on manually selected defect regions. If regions were segmented based on automated segmentation methods, it is possible that more variability would exist and, hence, lower classification accuracy could occur. Also, the defect classes were grouped based on anatomical and measure similarities, thus reducing variability within classes. The intent of this study was to remove the variability associated with segmentation, so that the interaction of scanning measures and classification rates could be analyzed for various defect types. For example, loose or unsound knots are complex features, which are a composite of knot, bark, checks, and other features. To minimize the effect of variability associated with such composite features, the knot class in this study was limited only to intergrown knots. Automated industrial inspection systems would have to fully understand and represent the variability of many more defect classes before they could be implemented effectively.

Another limitation of this study is the limited number of samples for each defect. Nineteen to twenty samples were collected for each defect type, and this number is relatively small for the quadratic discriminant functions used in comparing the classification accuracies of various sensor measure combinations.

It was determined that single sensor measures gave overall classification accuracies of less than 71%, indicating the need for multisensor and multimeasure methods for defect classification. It was discovered that the mean red color measure provided the highest overall classification accuracy (approximately 70%) between the defect classes studied. Adding information about color variability (e.g., the standard deviation of red color variation) helped further improve defect classification accuracy, especially where the differences between color measures were small in relation to the variation between these measures. Shape and density measures alone were found not to provide good classification variables for defects when used separately (e.g., less than 54% overall accuracy), but when used in combination with other measures contributed 5–10% to the classification accuracy between individual defects and 3–10% for the overall classification accuracy.

It was determined that the ANOVA method of model measure selection does not provide a direct method of selecting the optimum combination of measures; however, it was noted that the ANOVA results often do provide insight as to which measure should be selected in cases of confusion between two specific defect types. The knowledge gained about defect measure relationships was also valuable in explaining classification errors. By understanding how measures contribute to classification errors, future classification errors may be avoided. It was found that most classification errors were based on confusion with stain/ mineral streak. Most single measure classifiers performed 53% or lower in classification accuracy for stain/mineral streak. The ability to classify stain/mineral streak in multiple measure classifiers was 100%, proving that combining measures from multiple sensors increases classification accuracy of wood defects.

Certain feature measures were discovered to increase the accuracy in the classification of a particular defect type in a multimeasure discriminant classifier. An example of this is how the shape measure greatly increases the classification of stain/mineral and knots by 10–15%. The addition of a density measure was found to increase the accuracy of the final model from 94.9% to 98.8%.

It is likely that as variation increases in all measures, due to different segmentation methods or increasing the number of defect classes, that density and shape measures, and perhaps others, will become even more important for classification. It was determined that the measures taken from images used for defect classification for the defects included in this study should include two color mean values and two standard deviation values, a shape measure, and an X-ray standard deviation value. This information will help in the improved development of multisensor defect detection in the wood industry by providing information on what measures are best used in the classification of defects with color, shape, and density information.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the USDA National Research Initiative Competitive Grants Program through research grant number 9702090 and by the USDA Forest Service, Southern Research Station.

REFERENCES

Adel, M., D. Wolf, R. Vogrig, and R. Husson. 1993. Evaluation of colour spaces in computer vision application of wood defects detection. Pages 499–504 *in* Proc. IEEE International Conference on Systems, Man, and Cybernetics. v 2. IEEE Service Center, Piscataway, NJ.

Araman, P. A., D. L. Schmoldt, T. Cho, R. W. Conners, and D. E. Kline. 1992. Machine vision systems for processing hardwood lumber and logs. AI Applications 6(2):13–26.

Åstrand, E. 1992. Detection and classification of surface defects in wood: A preliminary study. LiTH-ISY-I-1437. Department of Electrical Engineering, Linköping University, Sweden.

Ballard, D. H., and C. M. Brown. 1982. Computer vision. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Bond, B. H. 1998. Characterization of defects in wood using color, shape, and density parameters. Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Brunner, C. C., G. B. Shaw, D. A. Butler, and J. W. Funck. 1990. Using color in machine vision systems for wood processing. Wood Fiber Sci. 22(4):413–428.

Buehlmann, U., J. K. Wiedenbeck, and D. E. Kline. 1997. Character-marked furniture potential for lumber yield increase in rip-first rough mills. Forest Prod. J. In Review.

Cho, T. H. 1991. A knowledge-based machine vision system for automated industrial web inspection. Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Conners, R. W., C. W. McMillin, and C. N. Ng. 1985. The utility of color information in the location and identification of defects in surfaced hardwood lumber. Pages XVIII-1–XVIII-33 *in* Proc. 1st International Conference on Scanning Technology in Sawmilling, San Francisco, CA.

———, C. T. Ng, T. H. Drayer, J. G. Tront, D. E. Kline, and C. J. Gatchell. 1990. Computer vision system hardware system for automating rough mills of furniture plants. Pages 777–787 *in* Proc. SPIE, Applications of Artificial Intelligence, VIII. Orlando, FL.

———, T. H. Cho, C. T. Ng, T. H. Drayer, P. A. Araman, and R. L. Brisbin. 1992. A machine vision system for automatically grading hardwood lumber. Industrial Metrology 2(3/4):317–342.

———, D. E. Kline, P. A. Araman, and T. H. Drayer. 1997. Machine vision technology for the forest products industry. IEEE Computer Innovative Technology for Computer Professionals 30(7):43–48.

Duda, R. D., and P. E. Hart. 1973. Pattern classification and scene analysis. John Wiley and Sons, New York, NY.

Everitt, B. S., and G. Dunn. 1992. Applied multivariate data analysis. Oxford University Press, New York, NY.

Forrer, J. B., D. A. Butler, J. W. Funck, and C. C. Brunner. 1988. Image sweep-and-mark algorithms. Part 1. Basic algorithms. Forest Prod. J. 38(11/12):75–79.

Grum, F., and C. J. Bartleson. ed. 1980. Color measurement. Academic Press, New York, NY.

HAGMAN, P. O. G., AND S. GRUNDBERG. 1993. Multivariate image analysis methods to classify features on Scots pine: Evaluation of a multi-sensor approach. *In* R. Szymani, ed. 5th International Conference on Scanning Technology and Process Control for the Wood Products Industry, October 25–27, 1993, Atlanta, GA. 18 pp.

HAND, D. J. 1981. Discrimination and classification. John Wiley and Sons, New York, NY.

———. 1986. Recent advances in error estimation. Pattern Recognition Letters 4:335–346.

HUBERTY, C. J. 1994. Applied discriminant analysis. John Wiley and Sons, New York, NY.

KAUPPINEN, H., AND O. SILVEN. 1995. Development of color-vision-based solutions for lumber grading. Pages 203–214 *in* Proc. SPIE-The-International-Society-for-Optical-Engineering. v 2423. Society of Photo-Optical Instrumentation Engineers, Bellingham, WA.

KLINE, D. E., R. W. CONNERS, D. L. SCHMOLDT, P. A. ARAMAN, AND R. L. BRISBIN. 1993. A multiple sensor machine vision system for automatic hardwood feature detection. *In* R. Szymani, ed. 5th International Conference on Scanning Technology and Process Control for the Wood Industry. October 25–27, Atlanta, GA. 11 pp.

LAMPINEN, J., AND S. SMOLANDER. 1996. Self-organizing feature extraction in recognition of wood surface defects and color images. Int. J. Pattern Recognition and Artificial Intelligence 10(2):97–113.

MEDIA CYBERNETICS. 1995. Image Pro Plus Version 1.3 for Windows. Silver Springs, MD.

NATIONAL HARDWOOD LUMBER ASSOCIATION (NHLA). 1994. Rules for the measurement and inspection of hardwood and cypress. Memphis, TN.

OTT, L. 1988. An introduction to statistical methods and data analysis. 3rd ed. PWS-Kent Publishing Company, Boston, MA.

PÖLZLEITNER, W., AND G. SCHWINGSHAKL. 1990. Real-time classification of wooden boards. SPIE. High-speed Inspection Architectures, Barcoding, and Character Recognition 1384:38–49.

PORTALA, J. F., AND J. CICCOTELLI. 1992. Nondestructive testing techniques applied to wood scanning. Industrial Metrology 2(3/4):299–307.

SAS INSTITUTE. 1996. SAS Users Guide. Cary, N.C.

SILVÉN, S., AND H. KAUPPINEN. 1996. Recent developments in wood inspection. Int. Pattern Recognition and Artificial Intelligence 10(1):83–95.

WEEKS, A. R. JR. 1996. Fundamentals of electronic image processing. SPIE. Washington, DC.

WIEDENBECK, J., AND U. BUEHLMANN. 1995. Character-marked furniture: The yield picture. *In* Proc. 1995 Hardwood Research Symposium, National Hardwood Lumber Association. 13 pp.