ELSEVIER Computers and Electronics in Agriculture 16 (1997) 255-271

Computers and electronics in agriculture

Machine vision using artificial neural networks with local 3D neighborhoods

Daniel L. Schmoldt [a,*], Pei Li[b], A. Lynn Abbott [b]

[a] *Southern Forest Experiment Station, USDA Forest Service, Brooks Forest Products Center, Blacksburg, VA 24061-0503, USA*

[b] *Bradley Department of Electrical Engineering, Virginia Tech, Blacksburg, VA 24061-0111, USA*

Accepted 12 December 1996

Abstract

Several approaches have been reported previously to identify internal log defects automatically using computed tomography (CT) imagery. Most of these have been feasibility efforts and consequently have had several limitations: (1) reports of classification accuracy are largely subjective, not statistical; (2) there has been no attempt to achieve real-time operation; and (3) texture information has not been used for image segmentation, but has been limited to region labeling. Neural network classifiers based on local neighborhoods have the potential to greatly increase computational speed, can be implemented to incorporate textural features during segmentation, and can provide an objective assessment of classification performance. This paper describes a method in which a multilayer feed-forward network is used to perform pixel-by-pixel defect classification, After initial thresholding to separate wood from background and internal voids, the classifier labels each pixel of a CT slice using histogram-normalized values of pixels in a 3 x 3 x 3 window about the classified pixel. A post-processing step then removes some spurious pixel misclassifications. Our approach is able to identify bark, knots, decay, splits, and clear wood on CT images from several species of hardwoods. By using normalized pixel values as inputs to the classifier, the neural network is able to formulate and apply aggregate features, such as average and standard deviation, as well as texture-related features. With appropriate hardware, the method can operate in real time. This approach to machine vision also has implications for the analysis of 2D gray-scale images or 3D RGB images.

Keywords: Image processing; Image segmentation; CT scanning; Hardwood logs; Forest products

1. Introduction

Hardwoods are popular as materials for furniture and fine woodworking due to their rich, colorful wood and their distinctive grain patterns. Since visual appearance is a

* Corresponding author.

Elsevier Science B.V.
PII S0168-1699(97)00002-1

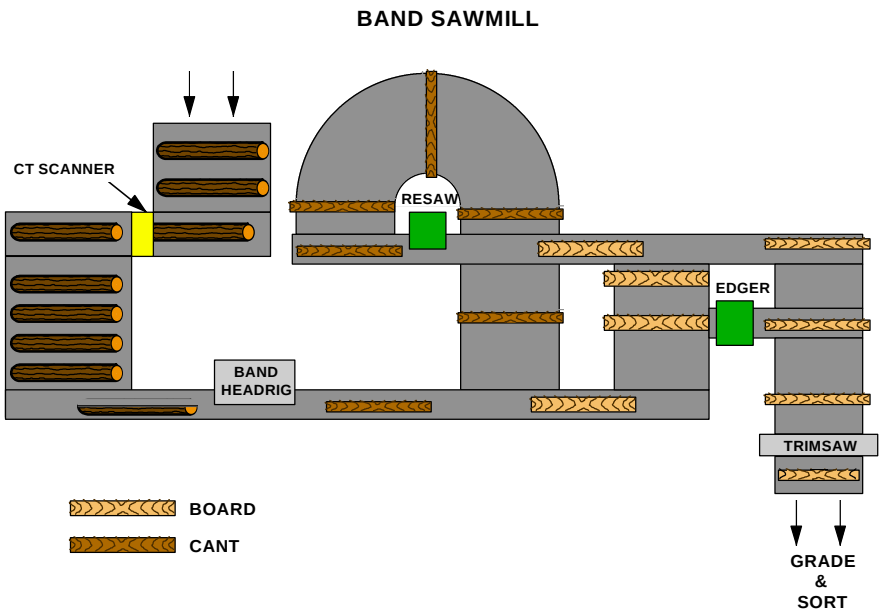

Fig. 1. A typical sawmill operation includes: a headrig for primary log breakdown, a resaw station for removing additional boards from the center cant of the log, and edger/trimmer stations for increasing the value of cut boards.

primary consideration, a “defect” is anything that adversely affects the wood’s aesthetic appearance. Although a large number of defect types have been cataloged, those of primary interest in this domain are knots, splits, decay, and bark. By properly sawing a log into lumber, many defects can be relegated to board edges or ends, where they can be easily removed. The careful choice of a sawing pattern leaves large, clear wood areas on each board, resulting in high commercial value.

In a typical sawmill, logs enter the mill and go through a de-barking process (Fig. 1). 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 the log interior is exposed with each board removed, the sawyer may re-orient the log periodically to cut from the best side. 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 value. The cant (the center section of the log) remaining from initial breakdown enters a resawing operation where additional boards are cut. These are also edged and trimmed.

Knowledge of internal log defects, obtained by scanning, is a critical component of efficiency improvements for future mills (Occeña 1991). Nevertheless, before computed tomography (CT) scanning or any other type of internal log scanning can be applied in industrial operations, there are several hurdles that must be overcome. First, there needs to be some way to automatically interpret scan information so that it can provide the saw operator with the information needed to make proper sawing decisions. A sequence of

X-ray tomographs cannot be synthesized readily into a three-dimensional (3D) mental model by human operators (Schmoldt et al., 1993). For the purposes of sawing the log cylinder into high-value boards, this means accurately locating, sizing, and labeling internal defects. Second, this defect recognition procedure must operate at real time speeds, so that scanning, image reconstruction, and image interpretation and display can be integrated into mill processing. Third, a 3D display of a log and its defects for the sawyer is only the first step toward real efficiency. Eventually, the sawyer must be guided by computer-analyzed suggestions for the best log breakdown sequence, or have the sawing completely controlled by computer processing (Occeña et al., 1995). The work described in this paper addresses the first two issues, automated scan interpretation and real-time defect recognition.

Because most log features of interest are internal, a non-destructive sensing technique is needed which can provide a 3D view of a log's interior. Several different sensing methods have been tried, including nuclear magnetic resonance (Chang et al., 1987, 1989), ultrasound (Han and Birkeland, 1992), and X-ray. Due to its efficiency, resolution, and widespread application in medicine, X-ray computed tomography has received extensive testing for roundwood applications (Benson-Cooper et al., 1982; McMillin, 1982; Taylor et al., 1984; Funt and Bryant, 1987; Som et al., 1992; Zhu et al., 1991d; Hagman and Grundberg, 1995). As noted above, however, 3D log images require computer analysis before they can be useful in an industrial setting.

Previous work on automatically labeling internal log features has established the feasibility of utilizing CT images for internal log characterization. These researchers have employed a variety of methods to, first, segment different regions of a CT image and, second, to interpret, or label, those segmented regions. Often, image segmentation methods are based on threshold values derived from image histograms (Taylor et al., 1984; Funt and Bryant, 1987; Zhu et al., 1991a; Som et al., 1992). Histogram-based thresholds can be determined either statically or dynamically. All pixels in an image are then labeled as belonging to one of several, unnamed groups. Contiguous regions of similar pixels are then given meaningful labels in an interpretation (classification) step. Texture-based techniques have been applied to region labeling (Funt and Bryant, 1987; Zhu et al., 1991 c). Knowledge-based classification (Zhu et al., 1991b; Zhu, 1993), shape examination (Funt and Bryant, 1987; Som et al., 1992), and morphological operations (Sore et al., 1992) have also been used to label regions. Hagman and Grundberg (1995) used normalized pixel values in a scaled, 8 x 16 window to label knot types on veneer slices using either a partial least squares classifier or an artificial neural network (ANN). While this approach is interesting, the methods employed were contrived in the sense that regions to be labeled were pre-selected and centered in the analysis window.

In most cases, image analysis has focused on a single two-dimensional (2D) CT slice, although neighboring slices have been used for 3D filtering during preprocessing steps (Zhu et al., 1991a), for multiple-image operations to detect knots (Sore et al., 1992), and for generating 3D objects (Zhu, 1993). One reason for the 2D emphasis is that 3D analysis requires substantially more computations than 2D analysis, when the third dimension is added without reducing the initial 2D neighborhood size. Another reason that 3D processing of CT images has not been popular in the past is that

spacing between adjacent CT slices has been relatively large. Processing complexity is compounded when slice separation distance is large compared to the pixel size on a single slice. In the past CT scanners produced pixels with large distances between slices, whereas today's scanners can generate images with pixel sizes of 1.0 x 1.0 x 1.0 mm^3 or less.

While previous efforts have demonstrated feasibility, they have some serious limitations. First, reports of defect labeling accuracy are often either anecdotal, based on success in a training set, or based on a single test set. 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 effectively used in real 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 of arbitrary complexity. Third, texture information, which is critical for human differentiation of regions in CT images (i.e., image segmentation), has been used for region labeling only. Automated recognition algorithms should exploit texture information for segmenting image regions also.

This paper presents an alternative to the above approaches that has been developed with these limitations in mind. In contrast to the previous global approaches that separate the tasks of segmentation and region labeling, our approach operates using local pixel neighborhoods primarily, and integrates segmentation and labeling into a single classification step. A feed-forward artificial neural network has been trained to accept CT values from a small 3D neighborhood about the target pixel, and then classifies each pixel as knot, split, bark, decay or clear wood. In order to accommodate different types of hardwoods, a histogram-based preprocessing step normalizes the CT density values prior to ANN classification. Morphological postprocessing is used to refine the shapes of detected image regions. These steps are described in the next section.

2. Methods

As shown in Fig. 2, an X-ray CT scanner produces image slices that capture many details of a log's internal structure. The slice shown here contains 256 x 256 elements, each corresponding to a volume of 2.5 x 2.5 x 2.5 mm^3. Examples of clear wood and hardwood defects are indicated in the figure. Because CT numbers are directly related to density, CT images vary dramatically for different species and by moisture content. Therefore, a log that is freshly cut will produce different CT values than one that has had time to dry.

The CT image vision system that has been developed here consists of three parts: (1) a preprocessing module; (2) a neural-net based segmentation and classification module; and (3) a post-processing module. The preprocessing 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 x 3 x 3 window about the classified pixel. Morphological operations are performed during post-processing to remove spurious misclassifications.

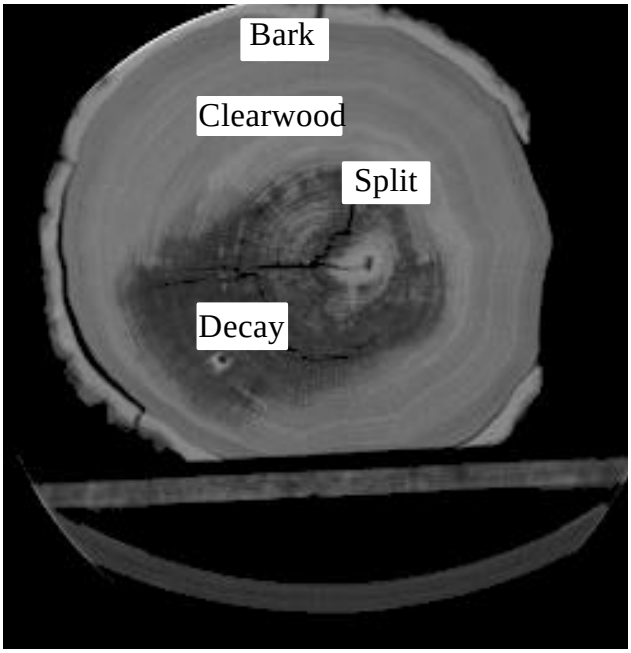

Fig. 2. Different densities are depicted by different gray-level values in this computer-generated X-ray tomograph of a red oak log cross-section. Regions of clear wood, decay, bark, and splits are visible. Each pixel is approximately 2.5 mm square.

2.1. Preprocessing

2.1.1. Background thresholding

The first objective of preprocessing is to identify background regions, so that these pixels can be ignored by the classifier. Our initial approach was to extract histograms for individual CT slices and apply Otsu's thresholding method (Otsu 1979). This method assumes bimodal histograms, and minimizes within-group variance. In our application, it automatically determines a correct threshold for many CT log images, because the histograms are typically bimodal. The two peaks can be found at very low gray-level values (background) and at relatively high CT values, corresponding to clear wood and high-density areas, such as knots and bark. Figure 3 illustrates this with a histogram of densities for the CT slice shown in Figure 2. In Figure 3, the rightmost histogram peak represents clear wood and bark. Knots are denser than clear wood, and tend to cluster at the right side of this peak when present. A large peak representing background is partially shown at the left.

Unfortunately, one of the defect types–decay–has density values which are roughly the average of background (air) and clear wood density values. This appears as a small peak in Figure 3, near the midpoint of the two larger peaks. If Otsu's method is applied directly to this histogram, the threshold indicated by t1 is detected. Unfortunately, this causes decay regions to be treated as background. We address this problem by weighting

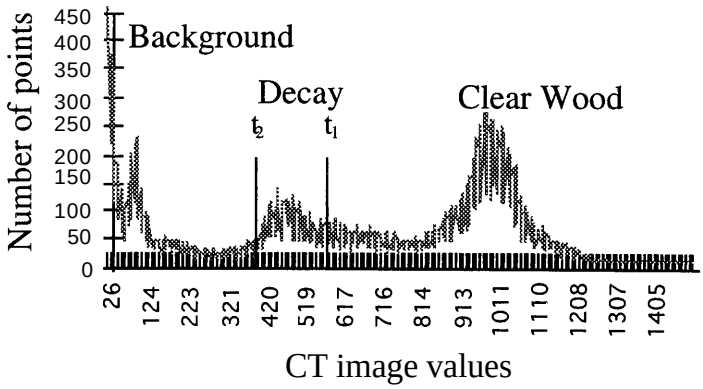

Fig. 3. Histogram of a log section. Background pixels produce a very large peak, part of which is omitted from the figure to improve clarity. The t_1 threshold is obtained using Otsu's method directly; t_2 is obtained after introducing a weighting function to the histogram.

the histogram values, using the function

$$w(t) = 1 - e^{-\left(\frac{t-t_1}{b}\right)^2} \quad (1)$$

where t_1 is the threshold determined by applying Otsu's method initially, and b = 2000. This value for b was chosen experimentally. The effect of weighting the histogram is essentially to remove the decay peak and reduce the size of the clear wood peak. If Otsu's method is applied to the resulting histogram, the threshold t_2 is found, which successfully distinguishes decay from background. This method has been tested using CT samples with both bimodal and trimodal histograms. The weighting function modifies histogram values only for the purpose of determining a threshold value for background pixels. The original pixel CT values are not modified in this step.

2.1.2. Density normalization

The second objective of preprocessing is to normalize CT values, so that the segmentation-classification step (hereafter referred to as just 'classifier') can work with different types of wood. Normalization is especially important because neighborhood pixel values are used as features by the classifier. If pixel values are not normalized there will be no consistent relationships among similar regions across CT images, and the ANN classifier will be unable to learn any useful patterns.

All hardwood CT histograms that we have examined have the characteristics of the histogram in Fig. 3. That is, there is a large peak of background pixel values at the far left, a large peak of clear wood, bark, and knot pixel values at the far right, and decay pixel values (if present) located at approximately the midpoint of the clear wood values.

To ensure consistency of defect region values across images, we need the ability to do several things with any histogram of CT density values. First, we want to shift the rightmost peak – containing clear wood, bark, and knot values – so that these regions always have the same values and so that the shape of this peak does not change. Second, we want the lower CT values, representing background, to remain about the same

following the transformation, that is, zero values stay near zero. Third, we want the CT values between the leftmost and rightmost peaks for each original histogram to have the same relative position in a transformed histogram. This type of transformation will give the important regions of any CT image the same density values, and allow us to apply our pixel-value dependent classifier to those normalized values.

The method used here applies a transformation to each CT value in the image. The transformation includes two components (1) a variable translation component, and (2) normalization by an arbitrary parameter. The transformation function is given in Eq. 2:

$$x_t = \frac{x_0 + f(x_0; x_{cw})(x_a - x_{cw})}{x_a} \quad (2)$$

where x_t is the transformed CT value; x_0 the original CT value; x_{cw} the original CT value of the clear wood peak; x_a the arbitrary translation anchor value, greater than the CT value of the clear wood peak; and f the translation multiplier.

The translation anchor x_a is an arbitrary parameter selected to be greater than the CT value of the clear wood peak. The rightmost histogram peak (including clear wood, knot, and bark values) will be shifted to the right by the amount $x_a - x_{cw}$, so that the clear wood peak is now at x_a. The resulting values are normalized by x_a so that the clear wood peak of a normalized histogram is always located at 1. In order for the translation of the rightmost peak to be consistent for all histograms it is necessary for the translation anchor value to be the same for all histograms. Otherwise, the shape of the rightmost peak will change with respect to the range of transformed density values.

The translation multiplier f is a function of the original CT value x_o and is parameterized by the clear wood peak value x_{cw}. It adjusts the amount of the maximum translation $x_a - x_{cw}$ that is added to the original value x_O to arrive at x_t after normalization by x_a. The actual equation for f is as shown below, Eq. 3. The function f is sigmoidal and symmetric about the value $x_{cw}/2$ (Fig. 4).

The range of f is $0 \leq f \leq 1$, where (1) the slope of f is very steep about the inflection point $x_{cw}/2$; (2) the value of f quickly approaches 0 at values of x_0 less than

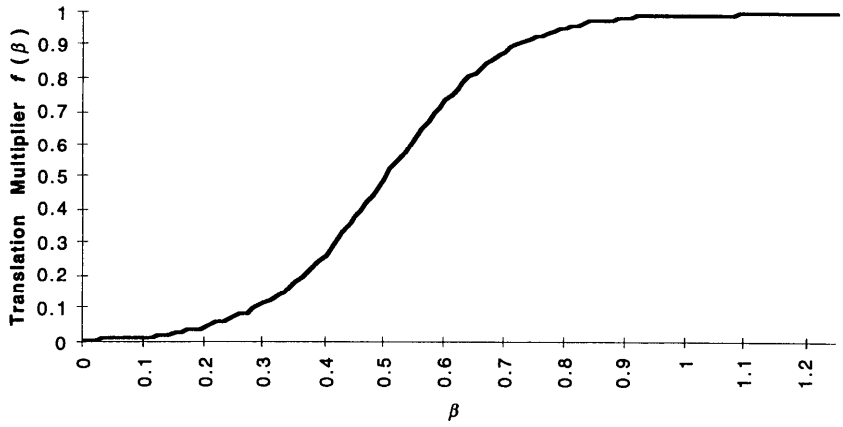

Fig. 4. The sigmoidal translation multiplier function $f(ß)$, where $ß$ is the proportion of the clear wood density value x_{cw}, adjust the amount of the maximum translation that is added to the original CT value.

$x_{cw}/2$; and (3) the value of f quickly approaches 1 at values of x_o greater than $x_{cw}/2$. At $x_o = x_{cw}/2$, f is exactly 1/2. The scale factor, adjusts the steepness of the curve about the inflection point, i.e., how quickly f rises from 0 to 1 as x_o increases. Larger values of α increase the steepness. Initially we have chosen $10/x_{cw}$ as a reasonable value for α.

$$f(x_0; x_{cw}) = 1/\left(1 + e^{\left(\frac{x_{cw}}{2} - x_0\right)\alpha}\right) \tag{3}$$

If we treat all CT values x_o as a proportion $ß$ of the clear wood peak value XC., i.e., $x_o = ß\, x_{CW}$ for some $ß$, then Eq. 3 can be rewritten as in Eq. 4, assuming = $10/x_{CW}$.

$$f(\beta) = 1/(1 + e^{5-10\beta}) \tag{4}$$

From Eqs. 2-4, we can observe that the following transformations will hold regardless of the original histogram:

$$x_t \cong 1 \quad \text{for } x_0 = x_{cw} \text{ or } \beta = 1$$

$$x_t \cong 0 \quad \text{for } x_0 = 0 \text{ or } \beta = 0$$

$$x_t \cong 0.5 \text{ for } x_0 = x_{cw}/2 \text{ or } \beta = 0.5$$

3. A neighborhood-based neural-net classifier

A multilayer feed-forward neural network is used to perform the primary segmentation-classification step. There were two initial goals in this research: (1) to determine if the heretofore separate tasks of image segmentation and region labeling could be combined into a single step, and (2) to determine whether an ANN classifier could perform well using only simple features obtained from small, local neighborhoods. Aside from initial background thresholding, both segmentation and defect labeling are performed simultaneously by the classifier. We have found that such a classifier works quite well, although performance is improved if information is also included concerning the distance of the target pixel from the center of the log slice. This distance measure provides contextual information that aids in classification, because some entities (such as splits) tend to lie near log centers and others (such as bark) lie near the outside edge of the log.

The classifier for a 3 x 3 x 3 CT window is shown in Fig. 5. As illustrated in the figure, each histogram-normalized value in the neighborhood serves as an input to the ANN. One additional input is the 'radius' of the element under consideration, which is the distance of this pixel "from the centroid of the foreground region of the CT slice. There are 5 output nodes of the ANN, one for each of the classes to be detected: knot, split, bark, decay or clear wood. The class associated with the output node that has the largest value for a given input is selected as the pixel label.

There are two types of split defects that are present in CT images and that we must treat differently in order to identify correctly. The first type of split is one that is wide enough to be imaged as an actual void, which then can be detected by background thresholding. The other type of split is a sub-resolution feature. It is visible in a CT image as a narrow, linear region of pixels with values near the low end of the clear wood values. These splits are narrower than the size of an image pixel, so when a pixel

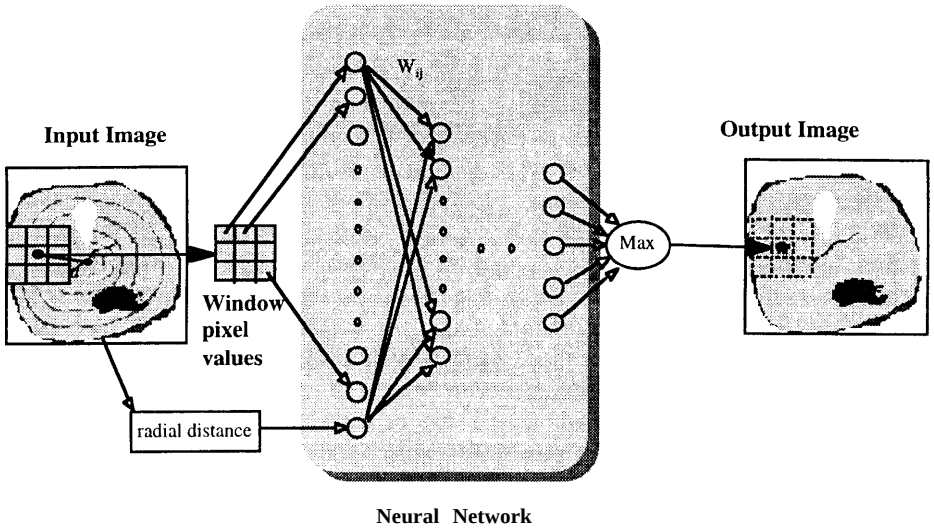

Fig. 5. Several different ANN topologies were trained using histogram normalized pixel values in a 3 x 3 x 3 window. The radial distance of the target pixel to the center of the log is included as a feature.

includes such a split, its CT value represents an average density for the void and the surrounding wood. The ANN classifier must be trained to recognize the texture pattern associated with such an anomaly in the clear wood region of an image.

The network was trained with the conventional back-propagation method (McClelland and Rumelhart, 1986) using Neural Works Professional II[1]. Because network topology has a large impact on classification accuracy and on convergence time during training, several topologies were compared. Networks using one, two, and three hidden layers were generated, with the total number of weights for each network topology kept constant (Nekovei and Sun, 1995; Özkan et al., 1993).

As of this date, the image interpretation system has been trained using only two hardwood species, northern red oak *(Quercus rubra,* L.) and water oak *(Quercus nigra,* L.). Although these two species are from the same family of oaks, they are from different geographic regions and growing conditions. Training/testing samples were selected from multiple CT slices. The entire training/testing set consists of 1973 samples. Ten-fold cross-validation was used to estimate the true accuracy rate of the ANN classifier.

In 10-fold cross-validation, the set of all samples is divided into 10 partitions. At each stage of the 10-step process, one of the partitions is reserved for testing, the classifier is trained on the remaining 9 partitions, and after training is complete the classifier is tested on the reserved partition. This process is repeated 10 times; final classification accuracy for the classifier is the average of the 10 test partitions. Cross-validation provides an objective and statistically valid estimate of the true classification rate (Weiss and Kulikowski, 1991).

[1]NeuralWare, Inc., Pittsburgh, PA. Trade names are used for informational purposes only. No endorsement by the U.S. Department of Agriculture or the Forest Service is implied.

3.1. Postprocessing

Because local neighborhoods are the primary source of classification features that are used by the ANN, spurious misclassifications tend to occur at isolated points. A post-processing procedure is used to remove small regions, thereby improving overall system performance. This method is effective because the defects of interest typically have relatively large sizes in an image.

We chose to use the gray-scale morphological operations of erosion followed by dilation for this purpose. A 3 x 3 structuring element is used for both operations. As the structuring element is moved from pixel to pixel across the image, each pixel value is replaced with the minimum (erosion) or maximum (dilation) value of the pixels in the 4-connected neighborhood defined by the structuring element. It was found that using 8-connectedness caused excessive erosion of split defects. Erosion removes isolated pixel values associated with small regions and smoothes the margins of larger regions. Dilation restores region pixels that were removed by the erosion operation. These two operations are typically combined to postprocess images with spurious pixel misclassifications (Gonzalez and Woods, 1992).

4. Results

Several sample histograms are presented in Fig. 6 to illustrate the effect of our density transformation procedure. Histogram appearance is invariant under this transformation, but the original CT values of critical regions have been automatically adjusted to be consistent across different CT images.

Four different ANN topologies were trained/tested using 10-fold cross-validation. The results are shown in Table 1. Classification accuracy is calculated as the proportion of test cases in which the correct class is the same as the output node with the highest value. Using this strategy, the ANN is either correct (1) or incorrect (0); classification is an all-or-none prediction. We also considered the results in terms of absolute accuracy, which is the average value of the output node corresponding to the correct class. For example, if *knot is* the correct class and the *knot* output node has the value 0.85, then the absolute accuracy for that test case is 0.85, not 1 as in classification accuracy. If, instead, the correct class for that test case is *decay* and the *decay* output node has the value 0.15, then the accuracy for that test case is 0.15. Using either classification accuracy

Table 1
Several neural network topologies were compared with respect to classification accuracy, absolute accuracy, and speed of training

Network topology	Number of weights	Number of training iterations	Absolute accuracy	Classification accuracy
28-12-5	396	6699	0.898275	0.947795
28-10-8-5	400	8299	0.902442	0.949316
28-7 -16-5	388	10499	0.869596	0.939686
28-8-8-8-5	392	60499	0.852903	0.853523

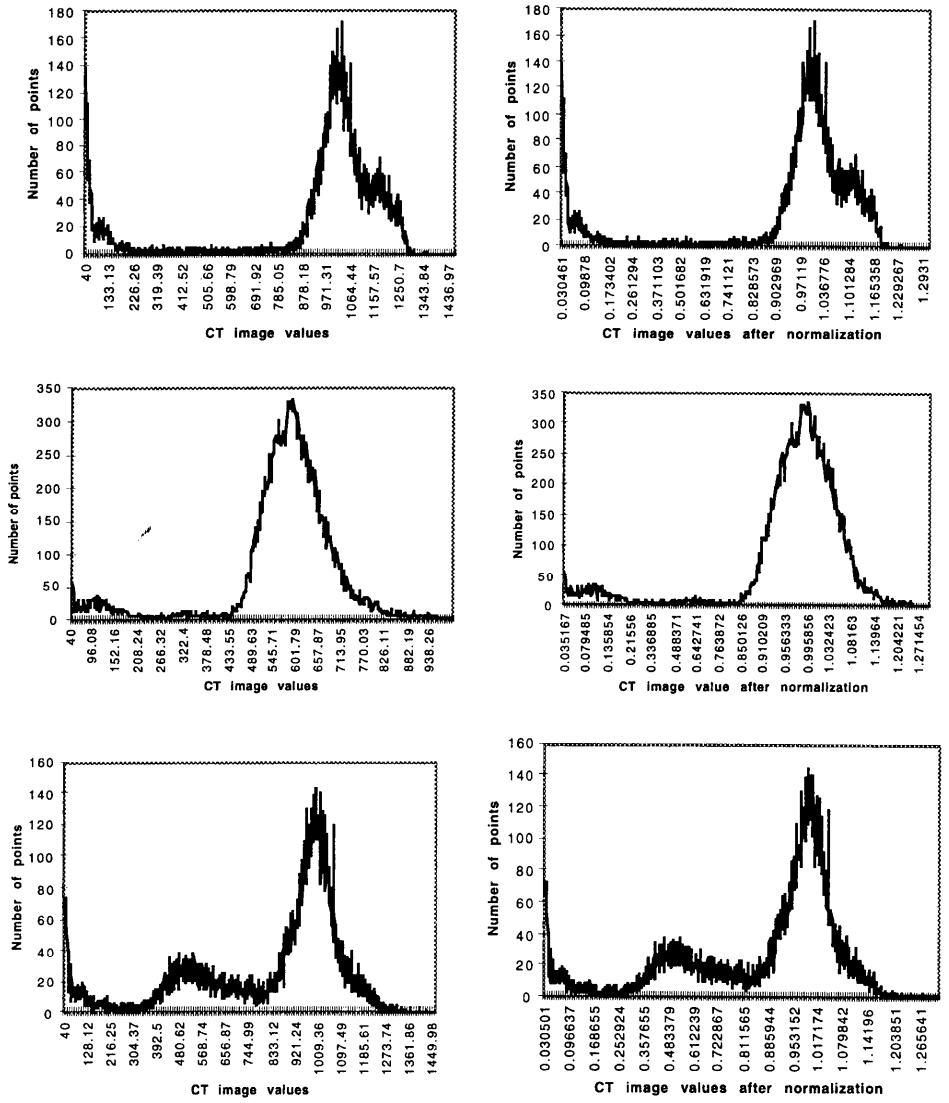

Fig. 6. Three CT image histograms illustrate the effect of transforming density values. Original CT image histograms appear on the left and transformed histograms appear on the right. Visually, the histograms do not change, while values for critical regions become approximately the same. The first and third histograms are red oak, and the second is yellow poplar.

or absolute accuracy, the ANN topologies compared in Table 1 have the same accuracy ranking.

The ANN with two hidden layers (28-10-8-5) exhibited the best performance with a classification accuracy of nearly 95%. The next best classifier (28-12-5), with a single hidden layer of 12 nodes, exhibited practically the same classification accuracy. Because

the latter network requires much less processing time, it was chosen as the optimal classifier among those evaluated. It is interesting to note that classification performance decreased slightly as the number of hidden layers increased. Experiments using different initial weights to train the networks indicated that the choice of initial weights has a negligible effect on the training process and on the performance of the classifier.

All of the different topologies exhibit high classification accuracy. There are much greater differences between them, however, when absolute accuracy values are compared. All of the classifiers are able to select the correct class in most cases (high classification accuracy), but the 2 top classifiers select the correct class using a much higher output node value, on average (high absolute accuracy). Although absolute accuracy values may eventually tell us something about the networks and how they will perform with other data sets, we are primarily concerned here with classification accuracy.

All of the neural networks considered here were trained by the delta rule with a momentum term. The delta rule modifies network weights in the training phase as a function of the global squared error. A learning coefficient (with a value between 0 and 1) moderates this weight adjustment, so that learning does not oscillate wildly back and forth but moves gradually in the direction of lower error. A momentum term acts like a sort of memory of past weight changes that counteracts the moderating effect of the learning coefficient. If a weight is continually moving in the same direction, the momentum term accelerates that movement by causing the squared error to increase the weight change. The effect of learning parameters on the speed of training convergence was studied by experimenting with various learning coefficients and momentum terms. Fig. 7 shows the experimental results. The final choice of the learning parameters is a small learning coefficient (0.1) and a medium momentum term (0.6).

Finally, we compared this 3D classifier with a similar ANN that used 2D neighborhoods only. Using only 9 pixels from a 2D neighborhood, rather than 27 pixels from the corresponding 3D neighborhood, classification accuracy dropped from 94.9%

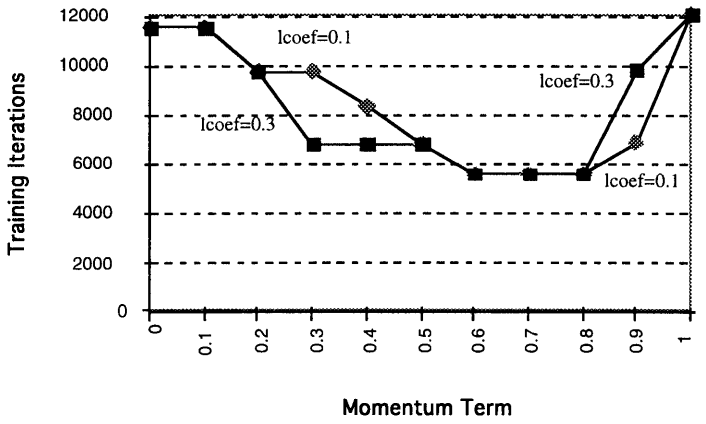

Fig. 7. Two different learning coefficients were compared with respect to training time for a range of momentum term values.

to 93.8%. This latter result is surprising in light of our on-going belief that there would be insufficient information present in a 3 x 3 neighborhood to produce a useful classifier.

The chosen classifier has been applied to several CT images for illustration. Four examples of processed log sections are shown in Fig. 8. The first 3 examples were chosen because they exhibit all of the defects of interest. The last example was chosen to demonstrate how the classifier performs on a species that it was not trained with, in this case yellow poplar (Liriodendron *tulipifera*, L.). As anticipated, the ANN produces some isolated pixel misclassifications, as shown in the middle column of the figure. The classification regions are improved with post-processing, however, as shown at the right. In the third example of Fig. 8, for example, the ANN classified partial regions of several growth rings as split defects; these were removed by subsequent postprocessing. In the upper two examples in that figure, incorrect labels near the outside border of the CT slices are removed by postprocessing steps.

Yellow poplar is very different in wood structure from oak. The classifier was not trained on any yellow poplar samples. Despite this, the classifier was able to distinguish bark and clear wood quite well. The knot area in the image is difficult to size correctly because it has CT values very similar to clear wood. It is not immediately clear whether we will be able to train the classifier to make this distinction, even by using yellow poplar samples.

This vision system is currently implemented on a Macintosh Quadra 650 containing an MC68040/33MHz processor. ANN classifiers were trained and tested using NeuralWorks. The best classifier was saved as C code, which was then incorporated into a XFCN code module that executes within an image processing software package (DIP-Station[2]). Our code module performs preprocessing and classification, and the image processing package does the postprocessing. Analysis of a single 256 x 256 CT slice requires about 25 s. This is considerably faster than the previous approach (Zhu, 1993) which requires 9 min of processing time on a VAX 11/785. Because the algorithms used in our vision system are implemented in C, they can be transported easily to faster computer hardware.

In comparison to previous hardwood log inspection systems, our system has a simple implementation, and at the same time high classification speed and accuracy. Other systems are reported to be able to successfully identify or locate some internal defects, but few statistical or quantitative results are available (Taylor et al., 1984; Funt and Bryant, 1987; Zhu et al., 1991a; Som et al., 1992; Zhu, 1993). Most previous work is limited to 2D image analysis (Funt and Bryant, 1987; Hagman and Grundberg, 1995; McMillin, 1982; Taylor et al., 1984; Zhu et al., 1991a), our results using both 2D and 3D data indicate that 2D may, in fact, be sufficient for CT images. Finally, most research has dealt with a single type of wood (Funt and Bryant, 1987; Hagman and Grundberg, 1995; McMillin, 1982; Som et al., 1992; Taylor et al., 1984), whereas our approach successfully deals with three different wood species.

[2] Digital Image Processing for the Macintosh, Perceptive Systems, Inc., Boulder, CO, USA.

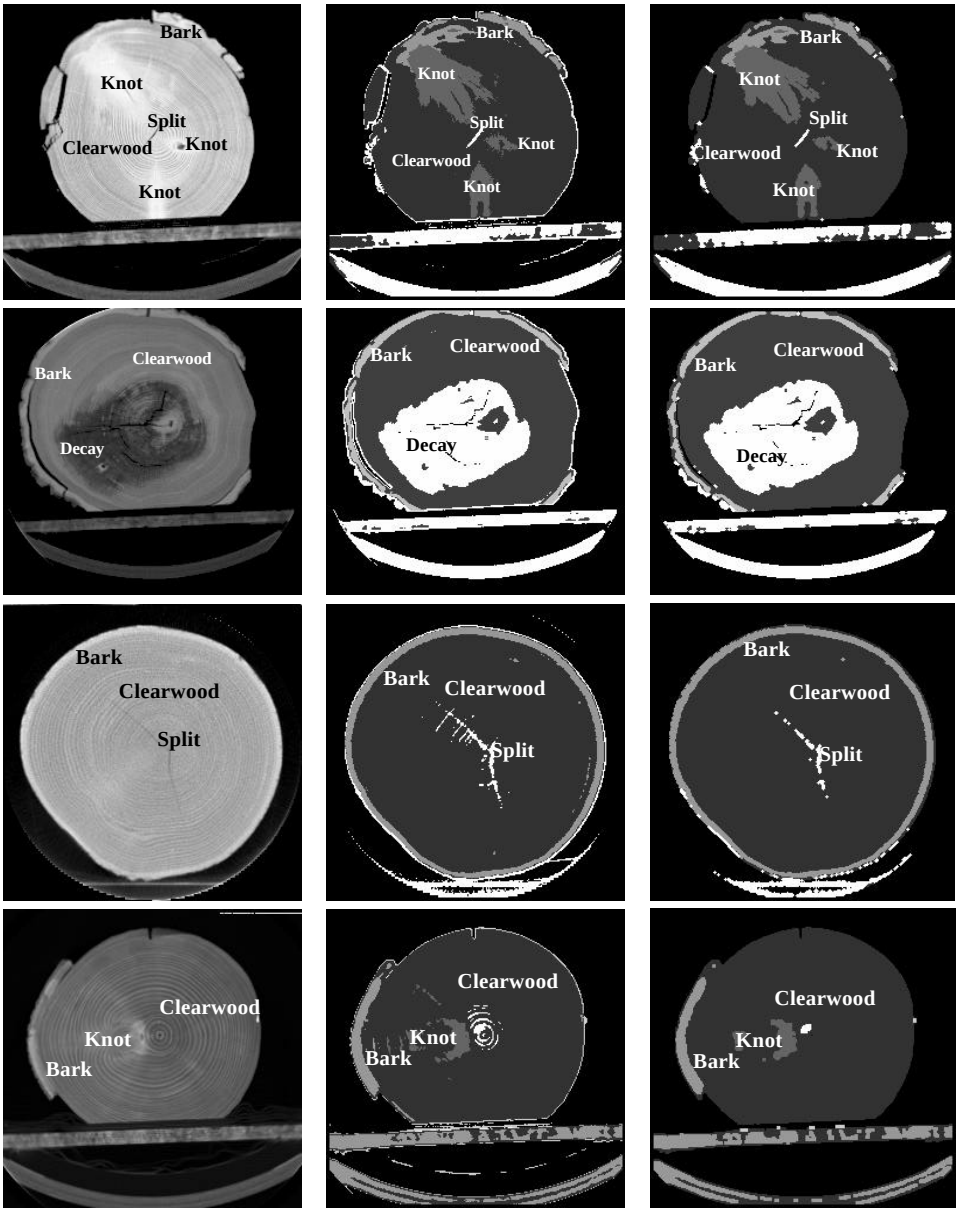

Fig. 8. Four log CT images demonstrate defect recognition results. Original CT images appear at the left in each row. The middle images are ANN classified images, and the rightmost images depict the classification results following postprocessing. The top 3 examples are oak and the bottom example is yellow poplar.

5. Conclusions

In most cases, the ANN classifier, operating primarily with local pixel values, is able to segment and classify regions of CT images with high accuracy. The resulting classification performance is 95% accuracy at the pixel level as confirmed by statistical cross-validation. Postprocessing improves this value considerably as confirmed by a visual examination, although we do not have a quantitative estimate for this improvement. Most regions are detected and correctly labeled; however, in some cases (e.g., yellow poplar knots) the classifier fails to correctly size defects. It is possible that by the addition of further postprocessing, e.g., high-level, rule-based analysis of defect region size and shape, we may be able to size defects more accurately and to remove any remaining misclassified regions.

It seems that the exact type of mapping (i.e., topology) between input and output does not have a large bearing on the results (at least, for this application). This makes sense both intuitively and mathematically. That is, for each of the 3D topologies, the number of weights was kept nearly constant. In a mathematical sense, this means that the number of parameters used in the mapping of input to output is the same for each topology. So, it stands to reason that each should work equally well, with only slight differences in the exact mathematical form of the mapping (resulting from the mathematical organization of the hidden nodes) accounting for the minor accuracy differences.

Of the two types of splits that are contained in CT images, the sub-resolution splits are difficult to correctly recognize. Image pixels containing the actual split have density values that lie within the clear wood portion of a CT histogram. Therefore, when the classifier recognizes these splits, it does so based on textural information rather than the mean value of pixels in the neighborhood, or a similar aggregate value. For other internal log features the ANN appears to use either texture information or a weighted average. For splits, however, it is highly doubtful that such an average could reliably discriminate sub-resolution features within the clear wood region of CT images. We conclude, therefore, that this ANN is truly a texture-based classifier for some features.

As noted above, the entire classification procedure requires only about 25 s on the current hardware. By using newer RISC-based hardware, this defect recognition time can be reduced drastically, by a factor of 8–10. This places defect recognition speed on a par with scanning and image reconstruction times. Because these 3 operations take 2–3 seconds each, they can be performed in parallel on successive slices. As scan i is being taken, scan $i-1$ can undergo reconstruction, and image $i-2$ from scan $i-2$ can undergo defect recognition. Therefore, this defect recognition technique can easily be implemented in real time as logs are scanned and images reconstructed.

Our preliminary test of the classifier on a species for which it was not trained has met with some success. Nevertheless, our only assessment of performance on an unfamiliar species has been visual and qualitative. We have not generated and applied a formal test set of samples similar to the original training set. Generally, bark and clear wood were classified correctly, however. Problems associated with misclassification of knot areas are due to the unique nature of yellow poplar knots. That is, these knots have CT values that are very similar to clear wood density values. Consequently, any defect recognition procedure that uses density values, e.g., CT data, will necessarily experience difficulty

with this species' knots. We plan to train a classifier specifically for yellow poplar knots with the hope that there are textural signatures unique to these knots that the ANN can learn.

6. Discussion

Although we have limited our investigations to 3D CT images of hardwood logs, it appears that the image analysis methods described here can extend to other applications and data types. Initial results using a 2D classifier produced only slightly lower classification accuracy than using 3D data. It is possible that a 2D classifier operating with a 5 x 5 neighborhood (a nearly identical number of inputs as in the 3D case) could perform equally as well, or better than the 3D-based classifier. Therefore, we feel that the same approach will work well with gray-scale video images, sonograms, and other 2D data types. There are other vision applications in which the spatial nature of the data is 2D, but the data's mathematical representation is 3D. This happens, for example, when color images are generated for a scene of interest. For applications involving color video images, it should be possible to treat the red, green, and blue (RGB) images as separate 'slices' which provide input to the ANN. Depending on the application, the ANN may produce final classification, or it may transmit information to a subsequent processing stage for higher-level analysis.

Because of the success of the trained ANN classifier on oak samples, we feel confident that we can develop species-dependent classifiers that are very accurate. We have plans to train species-dependent classifiers while at the same time retraining our generalized classifier for new species. Periodic comparisons of species-dependent versus species-independent classifiers should indicate whether species independence can be achieved. Should a generalized classifier prove to be infeasible, species-dependent classifiers can still be useful in actual mill operations because typically a single species is sawn over an extended production period.

Additional samples of CT images for other hardwood species need to be collected. This will enable us to verify the efficacy of our density normalization technique and the ability of the current classifier (or a newly trained classifier) to correctly label and size internal features of logs for other hardwood species. The ability to identify internal log features accurately and automatically is critical to the future adoption of internal log scanning in actual mill operations.

References

Benson-Cooper, D. M., Knowles, R. L., Thompson, F.J. and Cown, D.J. (1982) Computed tomographic scanning for the detection of defects within logs. Forest Research Institute, New Zealand Service, Bull., No. 8.

Chang, S.J., Olson, J.R. and Wang, P.C. (1989) NMR imaging of internal features in wood. Forest Products J., 39(l): 43-49.

Chang, S.J., Wang, P.C. and Olson J.R. (1987) Nuclear magnetic resonance imaging of hardwood logs. In: R. Szymani (Ed.), 2nd International Conference on Scanning Technology in Sawmilling, October 1–2, 1987, Oakland/Berkeley Hills CA. Forest Industries/World Wood, San Francisco CA. 8p.

Funt, B.V. and Bryant, E.C. (1987) Detection of internal log defects by automatic interpretation of computer tomography images. Forest Products J., 37(1): 56–62.

Gonzalez, R.C. and Woods, R.E. (1992) Digital Image Processing. Reading, MA, Addison-Wesley.

Hagman, P.O.G. and Grundberg, S.A. (1995) Classification of scots pine (Pinus sylvestris) knots in density images from CT scanned logs. Hols Roh- Werkstoff, 53: 75–81.

Han, W. and Birkeland, R. (1992) Ultrasonic scanning of logs. Industr. Metrol., 2(3/4): 253-282.

McClelland, J. and Rumelhart, D. (1986) Parallel Distributed Processing. Cambridge, MA, MIT Press.

McMillin, C. W. (1982) Applications of automatic image analysis to wood science. Wood Sci., 14(3): 97-105.

Nekovei, R. and Sun, Y. (1995) Back-propagation network and its configuration for blook vessel detection in angiograms. IEEE Transact. Neural Networks, 6(1): 64-72.

Occe-a, L.G. (1991) Computer integrated manufacturing issues related to the hardwood log sawmill. J. Forest Engin., 3(l):39-45.

Occeña, L. G., Chen, W. and Schmoldt, D.L. (1995) Procedures for geometric data reduction in solid log modelling. In: B. Schmeiser (Ed.), Proceedings of the 4th Industrial Engineering Research Conference, May 24-25, 1995, Nashville, TN.

Otsu, N. (1979) A threshold selection method from gray-level histograms. IEEE Transact. Systems Man Cybernet., SMC-9: 62-66.

Özkan, M., Dawant, B.M. and Maciunas, R.J. (1993) Neural-network-based segmentation of multi-modal medical images: a comparative and prospective study. IEEE Transact. Med. Imag., 12(3): 534-544.

Schmoldt, D. L., Zhu, D.P. and Conners, R.W. (1993) Nondestructive evaluation of hardwood logs using automated interpretation of CT images. In: D.O. Thompson and D.E. Chimenti (Eds.), Review of Progress in Quantitative Nondestructive Evaluation. Vol. 12. New York, Plenum Press, pp. 2257–2264.

Sore, S., Wells, P. and Davis, J. (1992) Automated feature extraction of wood from tomographic images. In: Second International Conference on Automation, Robotics and Computer Vision, September 15–18, Singapore. 5p.

Taylor, F. W., Wagner, J. F. G., McMillin, C. W., Morgan, I.L. and Hopkins, F.F. (1984) Locating knots by industrial tomography – a feasibility study. Forest Products J., 34(5): 42-46.

Zhu, D. (1993) A Feasibility Study on Using CT Image Analysis for Hardwood Log Inspection. Ph.D. Virginia Tech University.

Zhu, D., Beex, A.A. and Conners, R. W. (1991a) Stochastic field-based object recognition in computer vision. In: Stochastic and Neural Methods in Signal Processing, Image Processing, and Computer Vision, Vol. 1569, July 2 1–26, 1991, San Diego, CA. SPIE-The International Society for Optical Engineering, Bellingham WA, pp. 174-181.

Zhu, D., Conners, R.W. and Araman, P. (1991b) 3-D signal processing in a computer vision system. In: IEEE International Conference on Systems Engineering, August 1–3, Dayton OH, 4p.

Zhu, D., Conners, R. W., Lamb, F. M., Schmoldt, D.L. and Araman, P.A. (1991c) A computer vision system for locating and identifying internal log defects using CT imagery. In: R. Szymani (Ed.), 4th Internation Conference of Scanning Technology in Sawmilling, Oct. 28–31, San Francisco CA. Forest Industries/World Wood, San Francisco CA. 13p.

Zhu, D., Conners, R. W., Schmoldt, D.L. and Araman, P.A. (1991d) CT image sequence analysis for object recognition – a rule-based 3-d computer vision system. In: Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics, October 13-16, Charlottesville VA, pp. 173-178.