J Wood Sci (1999) 45:30–37

ORIGINAL ARTICLE

Jan Nyström · Olle Hagman

Real-time spectral classification of compression wood in *Picea abies*

Received: February 18, 1998 / Accepted: July 6, 1998

Abstract Compression wood is formed by the living tree to compensate for external loads. It creates wood fibers with properties undesirable in sawn products. Automatic detection of compression wood can lead to production advantages. A wood surface was scanned with a spectrometer, and compression wood was detected by analyzing the spectral composition of light reflected from the wood surface within the visible spectrum. Linear prediction models for compression wood in Norway spruce *(Picea abies)* were produced using multivariate analysis and regression methods. The resulting prediction coefficients were implemented in a scanning system using the MAPP2200 smart image sensor combined with an imaging spectrograph. This scanning system is capable of making a pixelwise classification of a wood surface in real time. Classification of one spruce plank was compared with analysis by scanning electron microscopy, showing that the automatic classification was correct in 11 of 14 cases.

Key words Compression wood · Imaging spectrometer · Smart sensor · Real-time classification · Multivariate modeling

Introduction

When a tree is exposed to abnormal loads during growth (e.g., a leaning tree), it starts to form reaction wood to compensate for the load. In softwood species, reaction wood is formed on the side of the stem exposed to compressive loads; hence the name compression wood. This is a common defect in Norway spruce *(Picea abies)*, which is the largest commercial species in Sweden. This investigation is limited to this species.

J. Nyström (✉) · O. Hagman
Luleå University of Technology, Division of Wood Technology, S-93187 Skellefteå, Sweden
Tel. +46-910-585300; Fax +46-910-585399
e-mail: jan.nystrom@tt.luth.se

Compression wood has a fiber structure that differs from that of normal wood. This affects the shrinking and swelling properties as moisture content changes, which is the cause of many problems involving compression wood. A review of what is known about compression wood can be found in Timell.[1]

Lumber containing compression wood tends to deviate from the intended shape after sawing and drying owing to the special fiber properties. Deformed wood may entail major handling problems that often lead to disturbances of production. These problems can be reduced by rejecting lumber that contains a large amount of compression wood at an early stage of the production line. An optical system with the capability to detect compression wood on wood surfaces could be used to reject the right pieces.

Compression wood has a chemical composition differing from that of normal wood. In particular, the lignin content is higher. Lignin has a characteristic absorption of light, especially high absorption of ultraviolet light.[1] The absorption of light in lignin from spruce has been analyzed by Norrström and Teder.[2] Using spectrophotometric methods, they showed that the spectral absorption consists of 13 absorption bands in the ultraviolet and visible spectra. These differences in light absorption can be used for automatic detection of compression wood by spectral analysis.

Spectral analysis produces an enormous amount of data. For real-time application it is essential to process the data close to the sensor to reduce the amount of data at an early stage. A smart sensor with built-in image processing electronics combined with an imaging spectrograph (spectrographic array with spatial information) makes up a scanning system that in real time is capable of making a pixelwise classification of the wood surface. The use of the prism grating prism (PGP) imaging spectrograph (Specim, Finland) combined with the MAPP2200 smart image sensor (Integrated Vision Products, Sweden) as a real-time classification system was presented by Åstrand et al.[3]

The system must be trained for the task. Multivariate image analysis methods were used to make linear models for different types of wood to distinguish compression wood from other normally occurring types of wood by its spectral characteristics. The linear models were implemented in the MAPP camera, which then was capable of making classifications in real time.

An evaluation plank was scanned, and the resulting classified image was compared to a human classification of the same surface. A number of samples were collected from the evaluation plank (primarily from areas with uncertain classification by humans or machine) and analyzed with scanning electron microscopy (SEM), which showed that the automatic classification was correct in most cases. The aim of this work was to develop prediction coefficients, based on spectral analysis of the wood surface, that are able to distinguish compression wood from normal wood in Norway spruce.

Materials and methods

All experimental work was done using dried and conditioned wood of Norway spruce *(Picea abies)* to avoid problems associated with keeping the wood surfaces in the same condition throughout the investigation. The scanning system described here is able to make classifications in real time, but it requires that the system has been trained to solve the problem. Figure 1 shows how the different stages of the training cycle are connected and how the same scanning system later is used for real-time classification.

Training set

A training set was put together that included different grades of compression wood as well as normal wood. The selected pieces of wood were considered to be good representatives of the entire material available for this investigation.

The training set consisted of eight wood pieces glued together, five of which contained more or less compression wood and three normal earlywood and latewood (Fig. 2). The pieces were oriented with both radial and tangential cross sections toward the surface, and the complete training set was sanded, forming a level training surface.

To verify which areas contain compression wood and which contain normal wood, samples from all eight parts were studied with SEM. Fiber properties such as circular cross section, thick cell walls, helical cavities, and intercellular spaces are typical for compression wood[1] and easy to establish with SEM.

Experimental setup

The PGP imaging spectrograph is arranged with the MAPP sensor as shown in Fig. 3. The PGP diverges light from a narrow line to a square sensor matrix, so very high intensity is needed on the line. For this reason a Solar 1000 sulfur plasma lamp[4] was used.

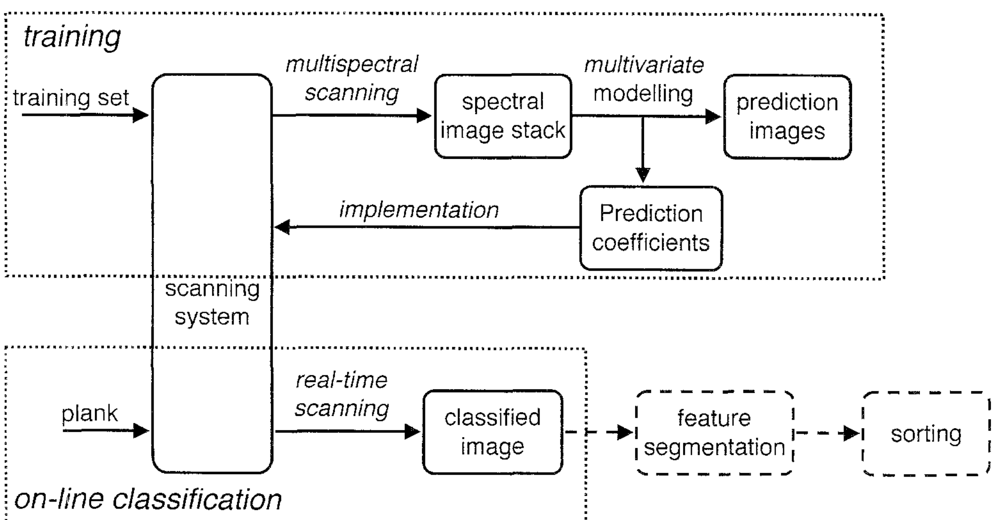
This electrodeless lamp uses microwave energy to heat sulfur, which emits light with a spectral distribution close to that of the sun and a high output that allows an intensity of 500000lux on the scene. Compared to a halogen lamp the sulfur plasma lamp emits little infrared radiation (heat). Radiation in the blue and near-ultraviolet region is also low, which is a disadvantage for this application because it narrows the effective spectral range.

The method of spectral analysis of the light reflected from the wood surface presupposes that the light penetrates the wood specimen and is colored by its light-absorbing characteristics. It is important to avoid specular reflectance from the wood surface. For this reason the MAPP–PGP system was arranged perpendicular to the wood surface and the illumination at an angle of 45 degrees (Fig. 4).

Multivariate modeling of images

Prediction coefficients for the MAPP–PGP system must be generated offline using a training procedure. The linear models for classification of compression wood were made

Fig. 1. Training cycle and real-time classification



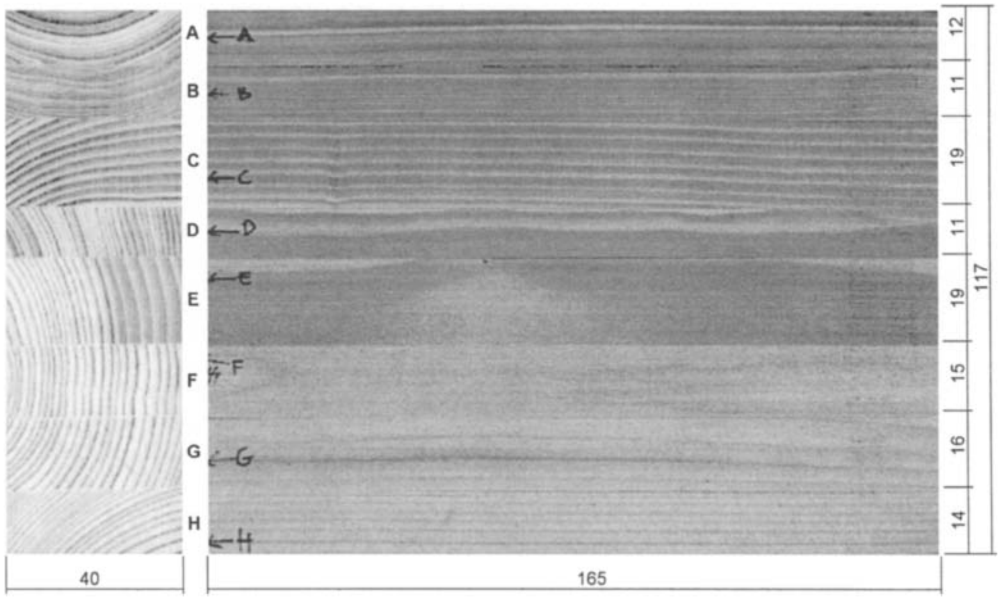


Fig. 2. Gray-scale image of the training set consisting of eight pieces of wood (*Picea abies*). Pieces *A–E* contain compression wood; *F–H* contain only normal earlywood and latewood

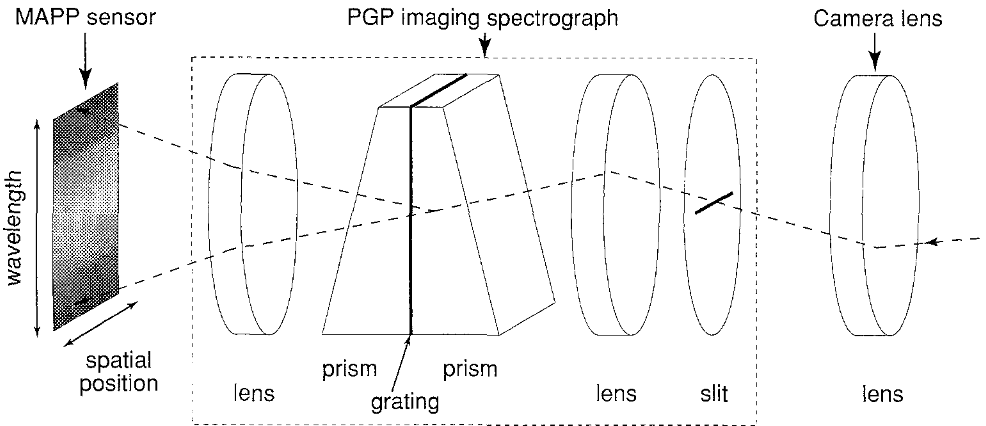


Fig. 3. Functional design of the imaging spectrometer. *PGP*, prism grating prism

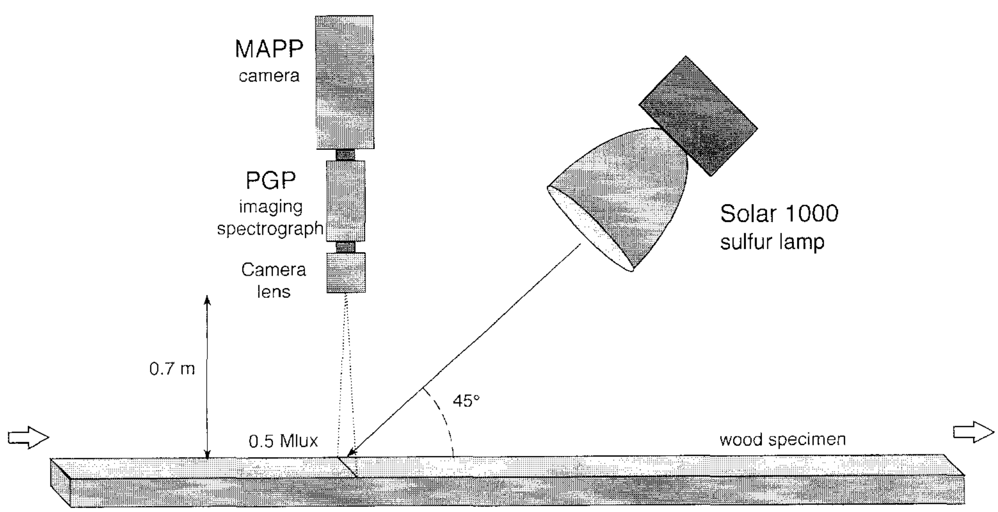


Fig. 4. Experimental setup

using multivariate analysis tools, primarily multivariate image analysis (MIA)[5] and multivariate image projections to latent structures (MIPLS)[6] implemented in the public domain *NIH Image* program (developed at the U.S. National Institutes of Health and available on the Internet at http://rsb.info.nih.gov/nih-image/).

Principal component analysis (PCA) is a method used for the analysis or compression of multivariate data by finding the dominating directions of the dataset in a multidimensional space. MIA is PCA applied to images. The result of MIA can be displayed as images, which makes interpretation of the result easier.

Partial least-squares regression, or projections to latent structures (PLS), is a method of iterative fitting of bilinear models in several blocks of variables and can make linear regression models for many classes simultaneously. MIPLS uses the kernel algorithm[7] for PLS, which allows many objects (pixels in the image) to be modeled toward several classes in a fast and memory-saving way. The result of MIPLS can be displayed as prediction images (prediction coefficients applied on the data set).

Generating the prediction coefficients

To generate the prediction coefficients needed for a real-time classification, the well-defined surface of the training set was scanned with the PGP–MAPP image spectrometer. Using the MAPP sensor as an ordinary camera results in 256 images of the same surface covering different spectral bands within the visible spectrum, each image with a spatial size of 256 × 64 pixels and an 8-bit gray scale. These images covered the entire training set, which means that the spatial resolutions were 0.45 mm/pixel crossboard and 2.5 mm/pixel in the grain direction. The images all represent different wavelength bands ranging from about 400 to 710 nm, giving a theoretical spectral resolution of 1.2 nm (Fig. 5). The image stack was then used as input data for the modeling. The prediction coefficients were generated by multivariate analysis primarily using MIA and MIPLS implemented in the image processing software *NIH Image*.

The entire data set was first analyzed with MIA to identify separable features and extract areas with a similar spectral pattern, which are considered good representatives of each class. By using score scatterplots to distinguish different features, good representatives for the training classes were obtained. In the next step the class representatives were input as a key with the original data and modeled by MIPLS. The resulting coefficients from MIPLS were then used as prediction coefficients in the PGP–MAPP system.

To give an instant appraisal of the prediction ability of the model, the prediction coefficients were applied on the spectral training data set, resulting in theoretical prediction images of the training surface (Fig. 6). These prediction images also indicated how many principal components (influential directions in a multivariate space) were needed to solve the problem. For a complex problem, more principal components are needed.

To distinguish compression wood from normal wood (noncompression wood), the training set was divided into four classes: compression wood, latewood, earlywood, and dark defects. Compression wood was wood cells with characteristic fiber properties (verified by SEM analysis), latewood was the darker part of the annual rings, earlywood the lighter parts of the annual rings, and dark defects the background or black knots with low reflection of light. The latter three formed the class "noncompression wood" but could not be modeled in one class because of their spectral differences.

The MAPP readout logic is restricted[8] such that every sensor row (spectral band) can be used for only one class. Consequently, if many types of wood are to be classified simultaneously, the sensor rows must be distributed among all classes. In this investigation we made linear models of four classes, so the sensor rows were distributed into groups with every fourth row in each class. For example compression wood uses rows 0, 4, 8, . . . , 252, and latewood uses rows 1, 5, 9, . . . , 253. The prediction coefficients (Fig. 5) generated by this training procedure were then implemented in the system by controlling the integration time in the MAPP sensor.

Real-time classification

The MAPP2200 smart image sensor is well suited for this application because of the unique possibilities of controlling the integration time individually for each sensor row and making an analogue columnwise summation of the sensor rows, which means considerable and fast reduction of data.[3]

Each row on the sensor corresponds to one spectral band. The integration time for each row was set proportional to the prediction coefficient corresponding to that spectral band.

As the data were read they were temporarily stored in an analogue register. The data rows were summed by reading data from several rows without resetting the analogue register. This columnwise summation of the charges in all included rows reduced the data to a single prediction row per class before the A/D conversion. Although it is possible to calculate negative prediction coefficients, it is not possible to implement them. Therefore, positive and negative sums were calculated[8] (both with positive integration time). The negative sum was later subtracted from the positive sum, making the class predictions.

During scanning, the prediction values for all classes were calculated one line at the time using the built-in image-processing electronics of the MAPP sensor. The data were then sent to a host computer where the prediction values for all classes were mutually compared and the classification was done. Each pixel was allotted to the class with the highest prediction value in that point, resulting in a classified image of the scanned surface. The level of each class prediction could be adjusted with an offset to set the classes to an appropriate level relative to each other. Because the image was processed line by line during the scanning, the classified image could be displayed within a split second after the last line was read.

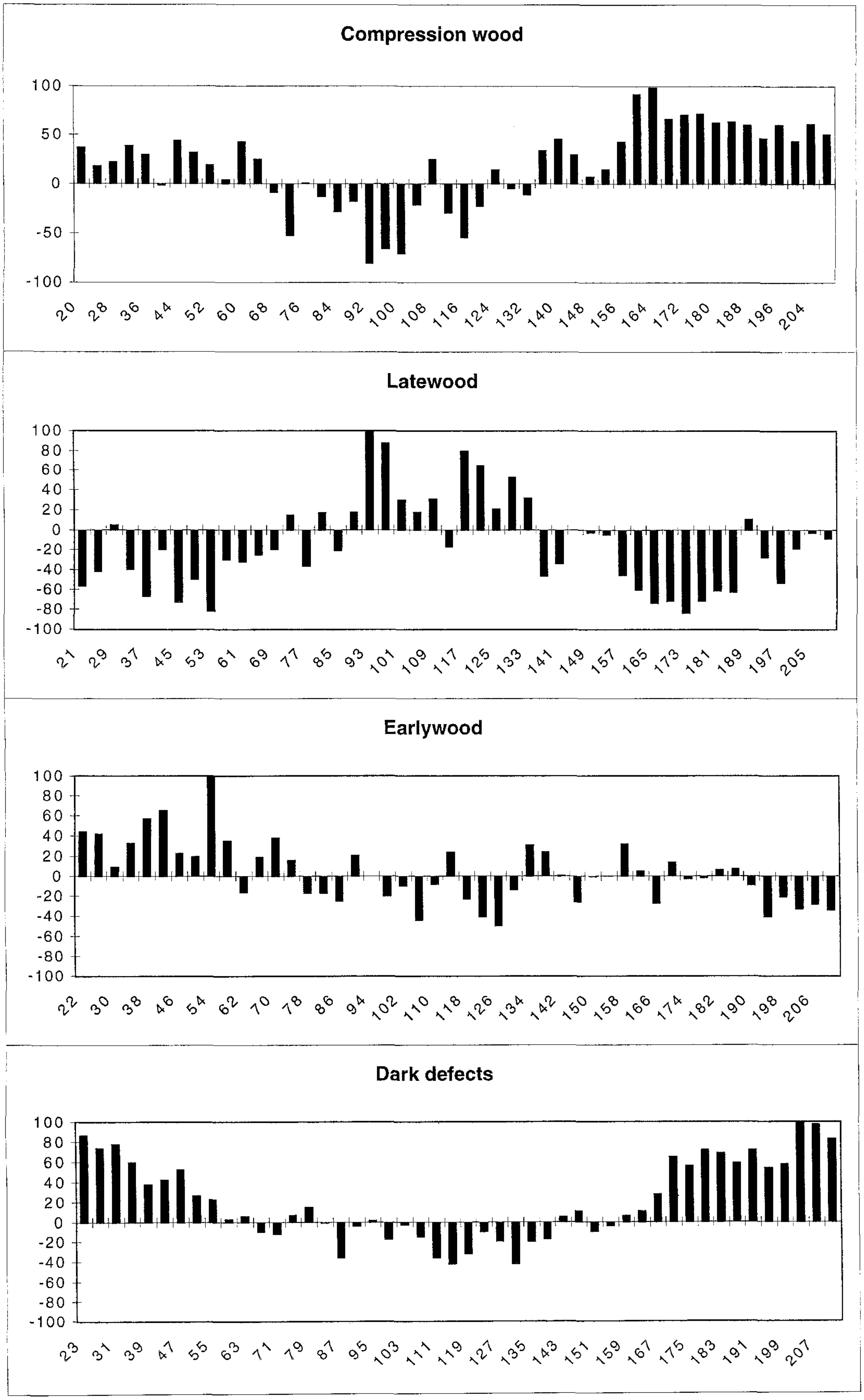


Fig. 5. Prediction coefficients modeled with multivariate image projections to latent structures (MIPLS). *Abscissa*: sensor row number, where rows 0 and 255 correspond to wavelengths of about 400 and 710 nm, respectively. *Ordinate*: percentage of maximum integration time

System evaluation

To evaluate the performance of the system, a plank of *Picea abies* was scanned and classified. The plank had a cross section of $125 \times 50\,mm^2$ and a length of about 4 m, and it was scanned with a speed of 0.3 m/s at a resolution of 2 mm/pixel in the feeding direction. The classification result was presented as a pseudocolored image where the colors represent different classes. This classified image was compared to the human judgment of the same surface but comparing only the presence or absence of compression wood. From the areas on the plank where the result from the automatic classification seemed to be incorrect or doubtful, samples were obtained and analyzed with SEM.

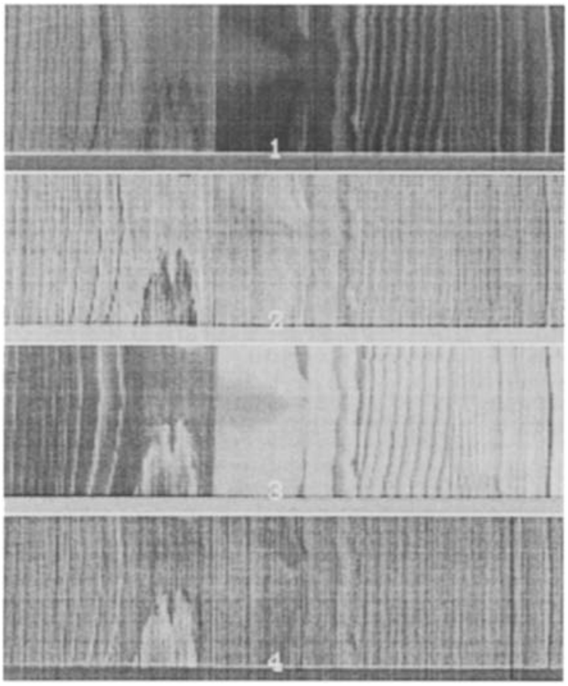


Fig. 6. Prediction images of the training set modeled with MIPLS. All images show the same surface with different prediction coefficients applied. *1*, compression wood; *2*, latewood; *3*, earlywood; *4*, dark defects. Gray-scale level indicates the probability of class possession; dark indicates high probability, light indicates low probability

Results

Training the system

Analyzing the training set with SEM verified that five of the eight pieces contained compression wood, mostly of rather severe grade in the latewood part of each annual ring, but in all cases together with a varying amount of normal earlywood. Three pieces contained only normal latewood and earlywood (Fig. 2).

The image stack of the training set was modeled by multivariate analysis, which generated four sets of prediction coefficients (Fig. 5). The best models were achieved with four principal components. Theoretical prediction images (Fig. 6) were generated at the same stage, giving a quick appraisal of the prediction ability of the model before the prediction coefficients were implemented in the scanning system.

Evaluation of real-time classification

The result from real-time classification of a plank (Fig. 7) was compared to a manual judgment of the same surface with respect to the presence of compression wood. Samples from the 14 selected areas on the plank were analyzed with SEM, which showed that the classification result could be considered correct in 11 cases and incorrect in three cases (Table 1). Scanning the plank in real time was done with a

Table 1. Judgment of the automatic classification compared to SEM analysis of the wood surface

Sample	Automatic classification	SEM analysis	Judgment
1	Thin streak of compression wood	Clear signs of compression wood in a streak of the annual ring	Correct
2	Scattered pixels of compression wood	Compression wood cells of less pronounced type in the middle of the annual ring	Correct
3	Scattered pixels of compression wood	Compression wood cells of less pronounced type are present in some areas.	Correct
4	Definitely compression wood	Clear compression wood	Correct
5	Compression wood	No signs of compression wood at the scanned surface, but further away in the annual ring	Incorrect
6	Compression wood	Clear signs of compression wood, but not in the entire annual ring	Correct
7	Definitely compression wood	Clear compression wood	Correct
8	Compression wood	Compression wood of varying grade in parts of the annual ring	Correct
9	Definitely compression wood	All signs of compression wood	Correct
10	Compression wood	Compression wood is present in the lower side of the branch	Correct
11	No indication of compression wood	Compression wood cells are present together with large regions of normal cells	Incorrect
12	Compression wood	Clear compression wood	Correct
13	Scattered pixels of compression wood	Compression wood cells of less pronounced types are present	Correct
14	Compression wood	No signs of compression wood	Incorrect

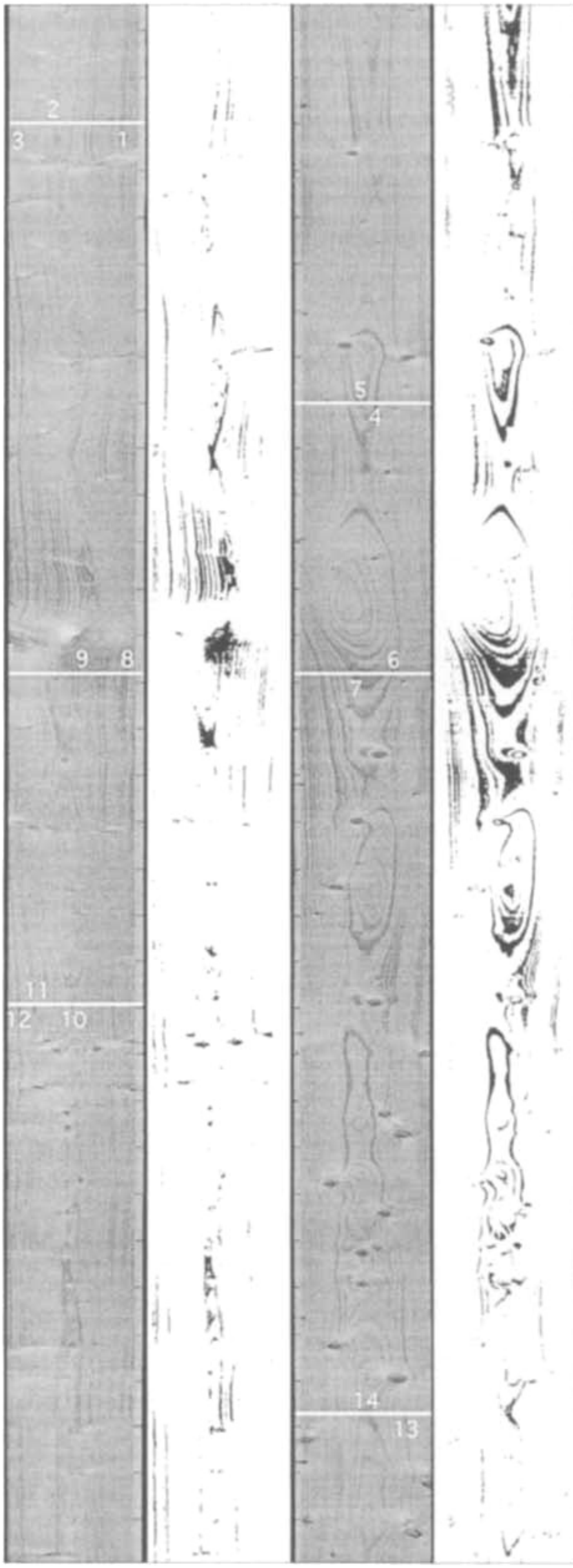


Fig. 7. Gray-scale images compared with compression wood classified images of the evaluation plank, pithwood, and sapwood side. Numbers denote samples analyzed with SEM. Black regions are compression wood in the classified images

sampling frequency of 150 lines per second with instant line-by-line classification.

Discussion

The best classification results were obtained with four principal components. If five or more principal components were used, the classification became worse, most likely due to overfitting (the noise in the data set was modeled).

Analyses of the samples with SEM show that the occurrence and grade of compression wood can vary considerably within each annual ring. That is why the manual judgment of the classification result is sometimes difficult. Generally, the classification seemed to be correct on areas with pronounced compression wood. Where the classification was uncertain, with scattered pixels of compression wood, analysis showed that these areas often contained unevenly occurring or a less pronounced form of compression wood.

When the classification was false, it was either on the edges of the plank where the compression wood class was too weak, or in the center of the plank where the class was too strong, which caused normal latewood to be classified as compression wood. This problem was probably due to unevenly distributed light on the surface (stronger light in the center of the plank). The best way to eliminate the problem is to change the setup of the lighting, but it is difficult to distribute light from a high-intensity point source evenly on a line.

All of the optical components have spectral characteristics that reduce the level of radiation in the blue/ultraviolet region, and many components limit the radiation in the red/near-infrared region. This narrows the spectral range, which is a severe limitation for the system, especially the lack of signal in the blue/ultraviolet region which is useful for detecting compression wood.

In this investigation only dried and conditioned wood was used to make sure the wood surfaces were kept in the same condition throughout the study, but it is of interest for an industrial application to make this classification on green lumber. Some kind of penetrating radiation, such as X-rays or microwaves, to obtain some estimation of the compression wood volume inside the wood could also be useful for predicting the form stability after drying.

The MAPP2200 sensor is controlled by a host computer that feeds the sensor with instructions. The instructions cannot overlap in time. Using the system for real-time classification almost certainly results in conflicts when finding time slots to execute readouts, sensor resets, and A/D conversions.[8] Running the software under Windows 95 probably adds distortion to the timing. If the instruction timing can be improved, the classification result is expected to improve as well.

References

1. Timell TE (1986) Compression Wood in Gymnosperms. Springer, pp 1–706
2. Norrström H, Teder A (1971) Absorption bands in electronic spectra of lignins. Svensk Papperstidning 74:85–93, 337–344
3. Åstrand E, Österberg M, Hagman O, Åström A (1995) Near-sensor real time spectral classification for industrial applications. In: Proceedings of the 9th Scandinavian conference on image analysis, Uppsala, Sweden, June 6–9, vol 1, pp 33–40
4. Anonymous (1997) Product Specification Solar 1000. IKL AB
5. Espensen KH, Geladi P (1989) Strategy of multivariate image analysis (MIA). Chemometrics Intelli Lab Syst 7:67–86
6. Hagman O (1996) On reflections of wood. PhD thesis, Luleå University of Technology
7. Lindgren F, Geladi P, Wold S (1993) The Kernel algorithm for PLS. J Chemometrics 7:45–59
8. Åstrand E (1996) Automatic inspection of sawn wood. PhD thesis, Linköping University