769. Wood products automatic identification based on fingerprint method

Ahti Põlder, Märt Juurma, Mart Tamre
Tallinn University of Technology (TUT), Dept. of Mechatronics, Ehitajate tee 5, Estonia
E-mail: *ahti.polder@ttu.ee*

(Received 10 September 2011; accepted 14 May 2012)

Abstract. This paper introduces utilization of improved fingerprint method based on wood surface visual parameters for increasing wooden products identifiability for instance in wood industry and especially for automation lines.

Usually the traceability of the objects in the production chain is achieved by using identification code based systems, for example visual codes, which are marked in one position and are read by a machine vision system in another position. In many cases the possible maximum readability can be quite hard to achieve due to marking problems or code background visual problems.

In wood industry traceability case it is possible to exploit the fact that wood is a biological material and therefore its surface is quite volatile, each code background is unique. The parameters describing the wood branches, grain or some other specific features can be used for matching for example the saw material in different production line positions. Presuming the conditions for the vision system in the consecutive positions are similar and no processing is applied to the traceable objects between those positions, then it is possible to match most of the objects and therefore increase automatic traceability and identification of the products in different positions of the production chain.

Keywords: identification, traceability, production automation, machine vision, fingerprint method, wood products, saw material.

Introduction

Monitoring of products and materials, tracing, inspection and respective control of the production processes gains increasing importance in today's production where the energy efficiency and reducing carbon footprint is of main importance alongside quality of the products. At consumer goods production and even for instance in metal working industry the principles are well adapted. In renewable materials industry like wood industry the principles are only gaining attention recently. Considering the wood production sector the target to implement the abovementioned principles is to investigate the quantitative measures of performance of the wood supply chain and quantification of the potential effects on the performance. Products and materials automatic or semi-automatic traceability should support finding tools and methods for holistic supply chain management, optimization and trade-off analysis by combining product quality and process economy with environmental impact from life cycle perspective, which could provide a new dimension to decision support systems in the industry and would avoid sub-optimization due to its holistic nature [1].

With this kind of traceability system the information loss can be avoided and decisions can be made according to knowledge, not only by experience. In order to gain information, a data acquisition infrastructure is needed and the data acquired has to be sensibly selected according to the business needs.

One important part of the wood supply chain is assuring the traceability of saw material. The saw material is marked with Data Matrix ECC200 barcode since it has good error correction and redundancy and the information density is much higher than traditional bar codes. The marked codes were read automatically in following production chain positions by vision system.

From the sawmill test runs it appeared that it is hard to achieve high code readability on automated lines and therefore material traceability over 95 % due to different problems regarding to material surface and marking quality is unachievable.

Since wood is a material characterized by an unique pattern of the surface the readability can be improved by implementing a specific technology utilizing additionally the fingerprint method, which is easy to automate in machine vision and which describes the objects specific parameters in parallel to reading the marked code. This approach is not as straightforward as identification code based system, but it can be used as a supplementation for this system. Fingerprint method is used in several different fields where biometrics can be used [2, 3].

In our paper [4] the board end surface histogram was used as a unique parameter. It appeared that histogram is quite sensitive to condition changes: brightness changes, board end size, defocusing and rotation. Lots of visually quite different board ends had similar histograms and that resulted in mismatches. One downside of histogram use is that the information gained has no position specific information, meaning that two different board ends can have similar histogram. Therefore the histogram based fingerprint method is in some respects acceptable but needs to be improved.

Board end *x* and *y* direction luminance linear averages

One way to improve the histogram based approach is to measure luminance linear averages of the board end. Linear averages on x and y axis direction are the arrays of mean luminance values of each pixel line in the image region of interest (ROI). The positions of different objects on the board end like code or branches are reflected on the directional averages making this method more unique than histogram only. To use directional averages for matching the board in different positions it is necessary to ensure that the board end is aligned horizontally. The algorithm was developed in LabView Professional Development System.

The measuring algorithm consists of image loading, board end finding and rotating. After that the linear averages, in *x* and *y* direction, are calculated of the rectangular ROI (Fig. 1.). From the averages plots it is possible to aim the location of the code. The application measures each image and no comparison takes place in this phase.

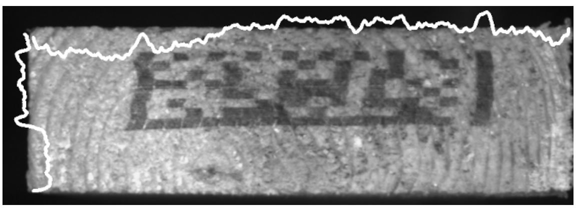

Fig. 1. *Y* axis averages (vertical plot) and *x* axis averages (horizontal plot)

The comparisons of measured parameters are made within another algorithm where linear correlation is used for comparing the *x* and *y* axis averages and histograms [5, 6]. In the current case, when comparing two images of the same board end from the different positions (or modifications), the correlation approaches to 1. Since linear correlation coefficient is independent of both origin and scale of the samples it is not necessary to rescale the measured values. That means a small unified change in lighting does not have effect on the comparison results.

To find the match of a certain board from one data sample it is necessary to find the correlation between that sample and all the samples from another data set where the data from images of the different position (modification) is. The highest of the resulting correlations is probably the match if the conditions are acceptable and the conditions in reading positions are similar.

Correlations between images acquired in single position

To evaluate if it is possible to match the images in between the different reading positions it is first necessary to estimate in what range the correlations between unmatching images are. For that two test were made where images from one position were compared with each other.

In the first test 500 images of marked board ends, acquired from saw line, were measured and correlations between all the measurement results were calculated (124750). The results show that the *x* axis linear averages correlation coefficients are distributed around 0.5. None of the correlation coefficients are in interval range from 0.95 to 1 (highest correlation is 0.94) indicating that *x* axis linear averages are potentially good candidate for fingerprint matching. In the *y* axis case the coefficients are distributed around 0.6 and 355 coefficients (about 0.3 %) are in range from 0.95 to 1. The averages of *x* axis are calculated around 3 times smaller amount of data then *y* axis meaning that y axis averages are not that sensitive to small differences on the board ends.

In the histogram case the number of correlation coefficients that fall into the range is over 8000 (about 6 %). That indicates that lots of histograms are similar to each other. That again indicates that the histogram is not very good parameter for comparison. Still, histogram can be used for pre-checking or for supporting some other parameter.

By multiplying *x* and *y* axis linear averages correlations we can describe both axes at the same time. In this case histogram peak is around 0 to 0.5 and no close matches exist, non of the multiplied correlation coefficients fall in the range from 0.9 to 1 (highest is 0.874). Therefore this parameter seems to be a very good candidate for the comparison. When multiplying multiplied correlation of *x* and *y* axis with histograms correlation then the correlation distribution histogram is even narrower, indicating that this parameter can probably give better differences in between matching images and other images correlation.

In the next test same sequence were followed on 282 images of board ends without marked code (39621 comparisons). The results were similar to previous test indicating that same approach can be used with the unmarked boards. From *x* axis linear averages none of the coefficients fall in the range from 0.95 to 1, from *y* linear averages 111 ones fall in that range (again about 0.3 %). In histogram case the corresponding values are 2283 and about 6 %. On multiplied *x* and *y* averages correlation distribution none of the values falls in the range of 0.85 to 1 (maximum correlation coefficient is 0.8417) and when multiplying that value with histograms correlation coefficient the values fall in even narrower set of bins.

Based on those tests results it is possible to estimate in what range the correlation values of one position most frequently fall and that there is some space (in the range from 0.9 to 1) for matching the images.

Correlations between original and modified images

To evaluate how the changes on the board ends change the correlation between the original images and modified images (or images from different position) one random image was selected from the both image groups and several modifications were generated from it. The image size was decreased and increased by 5 % to 15 %. Image brightness was decreased and increased. Also the dot was painted on the board end and on another case left half of the image was darkened; those images were used to test how the changes in the image structure affect the correlation coefficient.

Similar test were run as described before and linear averages and histograms of images were measured. Then all the modified images measured parameters correlation coefficients with original image parameter were calculated.

In general the images which brightness is changed have highest multiplied correlation, depending on amount of change, when images with changed size. For example when looking the images with marked code, images with changed brightness of 1 level gave the highest multiplied correlation coefficient 0.998, change of 2 levels gave 0.996 and change of 3 levels was lower. Same trend responds to images without marked code. That means the slight change in image brightness has minimal effect on matching results. This only applies when the images brightness is changed evenly on the whole board end. When the half of the board end is darkened the multiplied correlation is lower (no codes: -0.016 and with codes: 0.697) because the correlation is nonlinear.

Changes in the image sizes gave worse results. On the marked boards it results multiplied correlation values in between 0.211 to 0.512. In the case of unmarked boards it had not so strong effect (0.685 to 0.949). When there is no code on the board end, *x* and *y* linear averages are more likely within a smaller range and when the data array size changes (board end size changes) the trend remains and correlation is still quite high. When there are some strong changes in linear averages values (in a case of marked code), the change in the array size has much stronger effect on the correlation.

In the case of images with added dot or half image darkened, the multiplied correlation dropped with images without code. *X* and *y* averages of images with code had already some strong value changes and therefore an additional change does not have so strong effect in correlation (added dot: 0.581 and half darkened: 0.697).

Therefore the correlation coefficients when comparing the image with the modified image fall in the range mainly above 0.8. Thereby if the images from a real processing line have similar changes it is possible to match them.

Interpolating comparable data

To improve the correlation between the original and modified image the effect of the image size has to be decreased. One way to improve matching results is to interpolate the measurement data before finding correlation coefficients. Since the data sets only needs to be comparable and are used for comparison the linear interpolation can be used. The smaller array is interpolated so that it has the same number of elements as major array and the first and last element value remains the same. When the differences between images confine only in changes of the board end size on the image, the shapes of both arrays are almost identical [5, 6].

On the figure (Fig. 2.) the thin graph indicates the resized images data array (left) which is smaller. On the right the smaller array is interpolated and it is has same amount of data points and it is easily comparable with the other one.

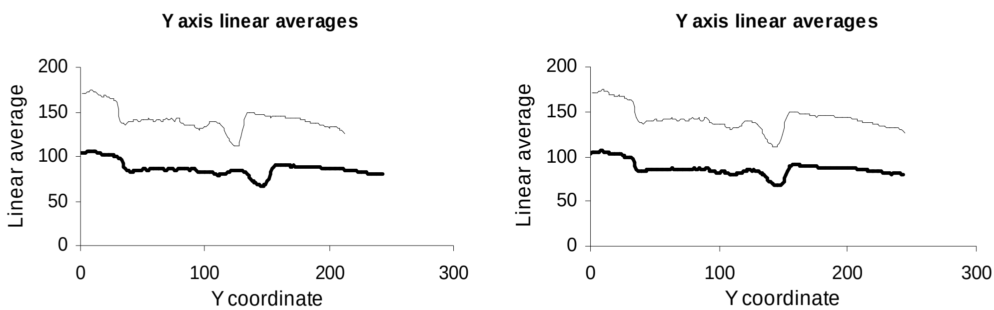

Fig. 2. *Y* axis linear averages of two images with code, original (left) and interpolated (right)

After implementing interpolation the same tests were run (one random image from both volumes of images, 13 modifications). It appeared that the interpolation had strong positive effect on correlation between the original image and its modifications. Most of modifications had correlation coefficient with original image over 0.9. Modifications with added dot and partial darken images were exceptional (since array size of those images were identical only the representation of pixels were changed) and basically did not change. Image with code modification which was brightest had correlation slightly below 0.9. With correct reading system setup this drastic brightness change is unreal. Those results are only a good base for following tests and learning how different setups can affect the results since the images were manipulated.

Final test

Finally two comparison tests were made. In the first test 13 modified images of an image with no code were added to the group of 282 and correlations between images were found. The highest multiplied correlation between images of different board ends was 0.83. The value of next multiplied correlation was 0.93. This and all the higher coefficients belonged to randomly selected image and its modifications. Some of the correlation coefficients in between the image modifications fall below 0.4 but in those cases the comparisons were made between images with half darkened or added dot. Therefore in from this test it is quite easy to match the images of the same board end.

In the second test same approach were used with marked images. On figure (Fig. 3.) left histogram indicates results without modified images and right histogram shows results with modified images. In this case the matches between the original and its modifications are more widely distributed; all values above 0.86 belong to randomly selected image and its modifications. The highest correlation between other not matching images is 0.85.

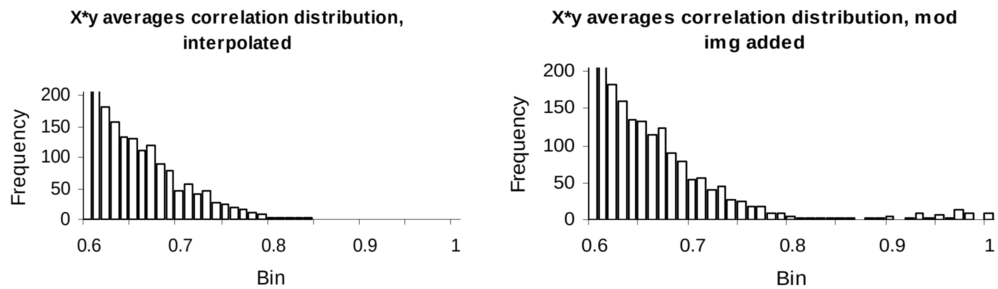

Fig. 3. Distribution of multiplied correlations without (left) and with (right) modified images

Therefore it is possible to match the images of board end modifications if the number of comparable objects is small. In real application that can be much more complicated since the number of images are probably larger and the differences in between comparable positions can vary more. In described tests the comparable images were modified manually and the actual differences in between the positions can be more complex.

Problems regarding to linear averages fingerprint method

Since reliability and readability (false matches are problematic) of described approach can vary and probably decrease proportionally to increase of comparable boards it can be used for matching relatively small amount of boards or as a support for the identification code based system for matching the boards which identification code was not readable.

It is necessary to link this system with code quality evaluation system. By doing that is possible to evaluate if it is necessary to measure the *x*, *y* linear averages. For example if the code was read in first position and the board end quality is good then the probability that this code is read in the following positions is higher than in a case when the quality parameters are low.

Other sorts of problems arise when something is changed on the board end. For example some additional dirt is applied or existing ice is melted or some process (like kiln) has affected the visual properties of board end. In this case described method probably does not give correlation good enough to make reliable matches and something more complex have to be used.

Therefore if the reading/measuring positions are correctly setup and the amount of comparable objects is small the probability of correct matches is high. On the other hand it is not tested on real production line and therefore can not be claimed.

Conclusion

The main scope of this paper is to give a proof of concept to simple fingerprint method based on linear averages for increasing the readability of the codes marked on the board ends which is an essential problem for wood production automation lines.

Several tests were run using images acquired from the real production line. Some random images were selected and modified by image manipulation software to simulate the possible changes on the images between different production line positions. In the real production line case this approach can have some drawbacks and therefore it needs to be tested in real production environment. In future a check method will be added to the system to eliminate the possibility that in a high correlation case the objects are actually not matching.

The histogram based comparison together with some board visual quality parameters can be used as a preliminary control, if they indicate the possible match then the linear averages are compared for verification. Therefore it is possible to classify the measured data and increase processing speed.

Acknowledgements

The first development of this approach originates from EU - 6th Framework Project Indisputable Key and is developed further at Tallinn University of Technology by support of the Grant SF0140113Bs08.

References

[1] **Uusijärvi R.** Annex I - Description of Work 2008-10-21: Detailed Implementation Plan for Next 18 Months for Month 19-36. http:\\ www.indisputable key.com, 2008, 154 p.

[2] **Chiorescu S., Grölund A., Berg P.** The finger print approach: using data generated by a 2-axis log scanner to accomplish traceability in the sawmill´s log yard. Forest Products Journal, Vol. 53, No. 2, 2003.

[3] **Maltoni D., Maio D., Jain A. K., Prabhakar A.** Handbook of Fingerprint Recognition. Springer, 2003.

[4] **Põlder A.** Modified fingerprint method for traceability enhancement of marked products. Proc. of 9th International Symposium „Topical Problems in the Field of Electrical and Power Engineering“, Tallinn: Estonian Society of Moritz Hermann Jacobi, 2010, p. 121 – 125.

[5] **Montgomery D. C., Runger G. C.** Applied Statistics and Probability for Engineers. 3rd ed., John Wiley & Sons, Inc., 2003, 174 p.

[6] **Terrell G. R.** Mathematical Statistics: A Unified Introduction. Springer, 1999, 76 p.