Optik 124 (2013) 949–952

Contents lists available at SciVerse ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Wood cell recognition using geodesic active contour and principal component analysis

Chen Guang-Sheng[a,b,*], Zhao Peng[a]

[a] Information and Computer Engineering College, Northeast Forestry University, Harbin City, 150040, China
[b] College of Materials Science and Engineering, Northeast Forestry University, Harbin City, 150040, China

ARTICLE INFO

Article history:
Received 17 October 2011
Accepted 15 February 2012

Keywords:
Wood cell recognition
Image processing
Principal component analysis
Shape

ABSTRACT

In this paper, we propose a robust wood cell recognition scheme using color wood cell images. First, a novel 2-D cell image collection system is devised, and the wood cell images are segmented by using a dual-threshold segmentation algorithm. Second, a geodesic active contour (GAC) is applied in the segmented binary image to extract the edge contours of multiple cells simultaneously. Third, wood cell recognition is performed based on the Mahalanobis distances calculated by using the principal component analysis (PCA) algorithm. We have experimentally proved that this scheme improves the recognition accuracy, which can efficiently discriminate the intraspecific cell's shape variation and the interspecific cell's shape variation.

1. Introduction

Wood cell detection is a significant issue in the wood industry [1]. It consists of the cell morphological analysis, cell density analysis, and wood cell recognition. A variety of visual characteristics has been used in the wood cell detection, and can be divided into three general categories: cell morphology [2–4], spectral analysis [5], and Fourier spectra [6,7]. As for the wood cell recognition, the enlarged cell images of the wood slice are usually picked up by the imaging camera. Then cell's shape features including the cell's area, perimeter, centroid, profile and number are extracted by the image processing algorithms for the subsequent wood cell recognition [8].

However, the current wood cell recognition based on shape features is sometimes poor especially for those cell species which have similar cell's morphological structures (e.g., for the two wood species *Abies nephrolepis* and *Picea jezoensis*, the cell's morphological structures are similar, as illustrated in Fig. 1). Moreover, for every wood species, the shape variations of cells usually occur in a cell image, as illustrated in Fig. 2. Therefore, the individual cell's shape variation consists of the intraspecific cell's shape variation and the interspecific cell's shape variation. The mixture of these two shape variations may decrease the cell recognition accuracy, since the intraspecific cell's shape variation may blanket the interspecific cell's shape variation. In fact, the current recognition schemes do not take into account these two shape variations with different natures.

In this paper, we propose a novel cell recognition scheme that considers the above-mentioned cell's intraspecific and interspecific shape variations. This recognition scheme consists of two advantages. First, it can efficiently discriminate the intraspecific cell's shape variation and the interspecific cell's shape variation by using the PCA algorithm. Second, a GAC model is applied in a binary cell image, which can extract the edge contours of multiple cells simultaneously in a parallel way. Therefore, this scheme is not only time efficient by using a GAC model but also improves the cell recognition accuracy by using the PCA algorithm which can both effectively tolerate the intraspecific cell's shape variation and effectively identify the interspecific cell's shape variation. We provide a fundamental experimental framework for a fast/accurate wood cell recognition work by using the cell's shape variation information. This framework can effectively and quickly identify the wood species by means of the cell recognition so as to judge the physical property and economic value of different wood species correctly. It can be used in some fields of the wood industry such as the wood assortment and the wood price so as to decrease the economic losses from the wood species misclassification.

2. Materials and methods

2.1. 2-D image collection system

One novel 2-D image collection system is devised as in Fig. 3. The image aiming system is the main part of an image collec-

* Corresponding author at: Information and Computer Engineering College, P. O. Box 319, Northeast Forestry University, Harbin City, 150040, China.
Tel.: +86 0451 82191532; fax: +86 0451 82191532.
E-mail address: kjc_chen@163.com (C. Guang-Sheng).

0030-4026/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijleo.2012.02.032

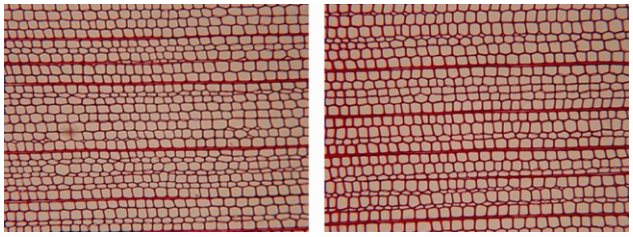

Fig. 1. Cell images of two wood species *Abies nephrolepis* and *Picea jezoensis*. (a) *Abies nephrolepis*; (b) *Picea jezoensis*.

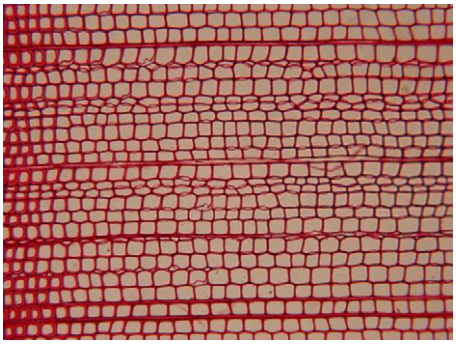

Fig. 2. Cell image of wood species *Picea koraiensis* to illustrate the intraspecific cell's shape variations.

tion system. In Fig. 3, the ray from the radian goes through the collimation path and is projected onto the detected object's surface (i.e., the wood microscopic slice). The cell image is formed in a stereo microscope and then picked up by a color CCD camera. After a photoelectric transform, this image is sent into the computer by the interface circuit to form the digital image which will be used in the cell recognition task. In our experimental setup, a XYH-3A stereo microscope and a Sony WV-CP240EXCH color camera are used to obtain the wood cell images. As for the accurate adjustment of the wood slice position in the X direction or Y direction, the optical scale can be used to fulfill this task, whose reading resolution is 0.1 μm here.

2.2. Image segmentation

During the production of wood microscopic slices, the wood slices are dyed by safranine. As a result of that, red is often the principal color channel in cell walls for every wood species, while no color exists in cell lumens. In our scheme, the RGB color information is combined into one color feature $(2R+G+B)/4$, since red is often the principal color channel in cell images. Therefore, the color cell images are converted into the grey-scale cell images with the combined color feature $(2R+G+B)/4$.

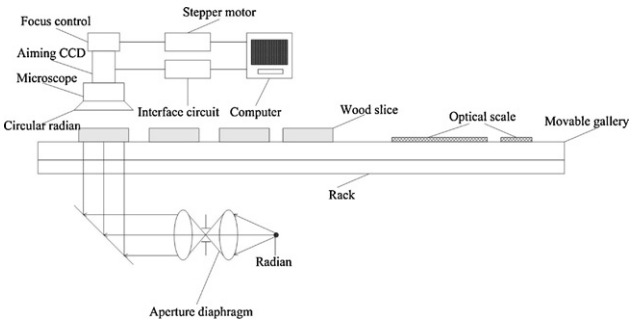

Fig. 3. 2-D cell image collection system graph.

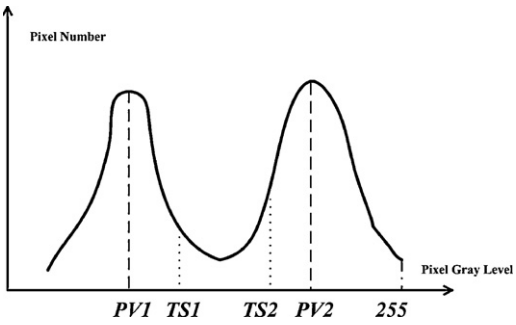

Fig. 4. The image histogram of the cell image.

For the converted grey-scale cell images, a dual-threshold segmentation algorithm is used:

$$I(x,y)=\begin{cases}0 & I(x,y)\le TS1\\ 1 & I(x,y)\ge TS2\\ \begin{cases}0 & \text{sum}\le 4\\ 1 & \text{sum}>4\end{cases} & \text{otherwise}\end{cases} \tag{1}$$

where $TS1$, $TS2$ are the dual thresholds of image $I(x,y)$:

$$\begin{cases}TS1=PV1+\frac{1}{5}(PV2-PV1)\\ TS2=PV2-\frac{1}{5}(PV2-PV1)\end{cases} \tag{2}$$

where $PV1$, $PV2$ are the 1st and 2nd peak values in the image histogram as illustrated in Fig. 4; $\text{sum}=\sum_{i=1}^{8}P_i$ is the statistical value in the 8-adjacent region of the processed pixel:

$$P_i=\begin{cases}0 & I(x,y)\le \dfrac{PV1+PV2}{2}\\ 1 & I(x,y)> \dfrac{PV1+PV2}{2}\end{cases} \tag{3}$$

2.3. Parallel contour extraction by GAC

Active contours are introduced by Kass et al. for segmenting objects in images using dynamic curves [9]. The existing active contours are classified as either parametric active contours or GACs. In particular, the parametric active contours are represented explicitly as parameterized curves in a Lagrangian framework, while the GACs are represented implicitly as level sets of the two-dimensional functional that evolve in an Eulerian framework.

GACs present one advantage over the traditional parametric active contours [10]. In fact, GACs represented by the level set functional may break or merge naturally during the evolution, and the topological changes are automatically handled. Therefore, they can detect multiple objects in an image simultaneously. In level set formulation, moving active contours are represented by the zero level set $C(t)=\{(x,y)\,|\,\varphi(t,x,y)=0\}$ of a level set functional $\varphi(t,x,y)$. The evolution equation of the level set functional φ can be written as follow [11]:

$$\frac{\partial\varphi}{\partial t}+F\left|\nabla\varphi\right|=0 \tag{4}$$

where the functional F is called the speed functional. For image segmentation, F depends on the image data and the level set functional φ.

Early GACs usually evolve the level set functional with a partial differential equation (PDE). Compared to pure PDE driven level set schemes, the variational schemes are more convenient and natural for incorporating additional information. In this work, we use a variational level set formulation of curve evolution proposed by Ming et al. [12]. For example, Fig. 5 shows the parallel contour extraction result by GAC in a binary segmented cell image obtained

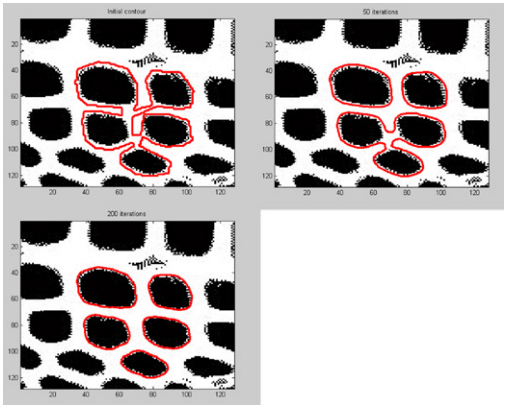

Fig. 5. Parallel contour extraction result by GAC in a binary segmented cell image. (a) Initial GAC; (b) result after 50 iterations; c result after 200 iterations.

by using the segmentation algorithm in Section 2.2. Moreover, every cell lumen's area (unit: pixel) is computed as follows.

(1) For a CCD image with multiple cells, we apply the variational GAC firstly. Then we find the background pixels with the following method: a pixel $p(x,y)$ is considered as one background pixel, if its level set functional value $\varphi(x,y) > 0$.
(2) Let's consider pixels as squares. The pixel $p(x,y)$ with its $\varphi(x,y) < 0$ and with all its four corner pixels (i.e., the upper, down, left and right pixels) having $\varphi < 0$ is considered as the cell's interior pixel, and then this pixel contributes one unit to the cell's pixel area integration.
(3) Let's find the cell's boundary pixels. The pixel $p(x,y)$ with its $\varphi(x,y) < 0$ and with at least one pixel having $\varphi > 0$ among its four corner pixels is considered as the cell's boundary pixel. The area computation in those boundary pixels is based on counting those pixel squares that are divided by the zero level set computed before.

After every cell lumen's area S_l is calculated, it is compared with two thresholds T_S^1, T_S^2 and those cells whose areas satisfy that $T_S^1 \le S_l \le T_S^2$ are retained to discard those small tissues and large adherent cells.

2.4. Shape recognition with PCA

2.4.1. Shape analysis of training cells

We use the GAC to simultaneously extract every cell's contour edge for each training cell image, species by species, to investigate characteristic shape variations of wood cells within the same species. Before the real analysis of shape variations could take place, it is necessary to align the set of training cell shapes. The alignment is achieved by scaling, rotating and translating the training shapes so that they correspond as closely as possible with each other. Each of the aligned training cell shapes gives rise to a vector $\mathbf{x}_i$ describing n boundary points:

$$\mathbf{x}_i = (x_{i,1}, y_{i,1} \ldots x_{i,n}, y_{i,n})^T \tag{5}$$

where $i = 1, 2, \ldots, N$, and $(x_{i,j}, y_{i,j})$ is the jth point of the ith cell shape; N is the number of training shapes for every species. The mean/standard cell shape vector $\bar{\mathbf{x}}$ is calculated as:

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{6}$$

The modes of shape variations, i.e. the ways in which the cell's contour edge points tend to move together, can be found by applying a PCA to the deviations from the mean vector:

$$\Delta\mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}} \tag{7}$$

From these deviations, the $2n \times 2n$ covariance matrix $\mathbf{S}$ can be computed as follow:

$$\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N}\Delta\mathbf{x}_i \Delta\mathbf{x}_i^T \tag{8}$$

The modes of variations of the cell shape points can be described by the $2n$ unit eigenvectors, $\mathbf{p}_1, \ldots, \mathbf{p}_{2n}$, and the corresponding $2n$ eigenvalues, $\lambda_1, \ldots, \lambda_{2n}$, of the covariance matrix $\mathbf{S}$. Any shape in the aligned training set can be approximated as a sum of the mean shape and a weighted sum of the first $t (t < 2n)$ eigenvectors:

$$\mathbf{x}_i \approx \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}_i \tag{9}$$

where $\mathbf{P} = (\mathbf{p}_1 \mathbf{p}_2 \ldots \mathbf{p}_t)$ is a matrix of the first $t (t < 2n)$ eigenvectors; and $\mathbf{b}_i = (b_{i,1}, \ldots, b_{i,t})^T$ is a vector of weights.

2.4.2. Recognition

Eq. (9) permits the generation of a new shape example by replacing $\mathbf{b}_i$ with a new vector of weight $\mathbf{b}$. Provided that the weights are not too far from zero, the new synthetic example will be similar to those in the training set. For an unknown cell shape vector $\mathbf{x}_l$ in a cell image of unknown species (i.e., $l = 1, 2, \ldots, K$; K is the cell number to be recognized), the weight vector $\mathbf{b}_l = (b_{l,1}, \ldots, b_{l,t})^T$ is computed as follows:

$$b_l = \mathbf{P}^T(\mathbf{x}_l - \bar{\mathbf{x}}) \tag{10}$$

Define the Mahalanobis distance D_l as follow:

$$D_l^2 = \sum_{h=1}^{t}\left(\frac{b_{l,h}^2}{\lambda_h}\right) \tag{11}$$

The final classification distance is:

$$D = \sum_{l=1}^{K} D_l \tag{12}$$

Therefore, an unknown cell image should be classified as a species to which the minimum distance $D_{\min}$ is obtained.

3. Results and discussion

3.1. Results

The wood cell images are acquired by the experimental platform as illustrated in Fig. 3. To process the cell image sequence efficiently, DH-CG400 video image collector is used to achieve a fast transmission speed of 132 MB/s. Experiments are performed in the Pentium 2.80 GHz computer with the internal memory of 1.0 GB by using the Matlab 6.5.

Five different wood cells including *A. nephrolepis*, *Picea koraiensis*, *Pinus koraiensis*, *Cunninghamia lanceolata*, *P. jezoensis* cells are tested, as illustrated in Fig. 6. For every wood cell species, 30 images are picked up by the camera. Among these 30 cell images for every species, 15 images are selected as the training cell images. Then the pattern recognition experiment is performed to the remaining 15 images and the relevant recognition rate of accuracy (RRA, i.e., the percentage rate of correctly identified wood cells) is calculated. The relevant RRAs for five wood species *A. nephrolepis*, *Picea koraiensis*, *Pinus koraiensis*, *C. lanceolata*, *P. jezoensis* are approximately 85%, 83%, 89%, 80% and 91%, respectively.

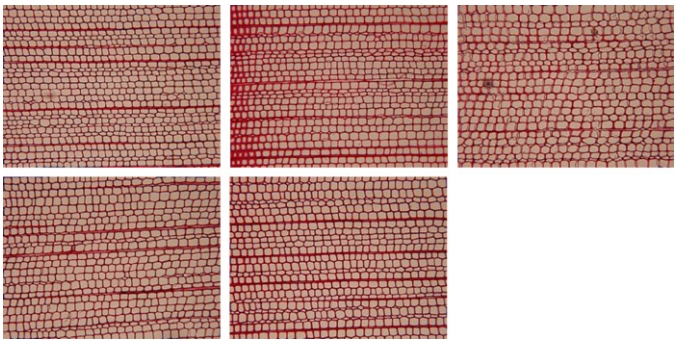

Fig. 6. Cell images of five wood species. (a) *Abies nephrolepis* cell; (b) *Picea koraiensis* cell; (c) *Pinus koraiensis* cell; (d) *Cunninghamia lanceolata* cell; (e) *Picea jezoensis* cell.

3.2. Discussion

To accelerate the recognition speed and improve the recognition accuracy, two issues should be remarked. First, the original cell image size is large (i.e., with a resolution 1392×1040) and there are hundreds of cells in one image. Obviously, it is not practical/possible for a GAC to converge to the contour edges of all these cells simultaneously. Therefore, we segment the original cell image into many small sub-images (i.e., with a resolution 128×128). In these sub-images, there are usually 10 wood cells approximately and our GAC will often converge to the contour edges of the interior 5 cells simultaneously. The following cell recognition work is performed in these cell sub-images. For each sub-image, a classification distance D is computed and the final classification distance for the original cell image is the mean value of all classification distances in sub-images.

Second, to ensure the contour extraction accuracy of the GAC, the initial GAC must be placed near the contour edges of the detected cells manually; otherwise, the GAC may converge to the wrong contour edges. Some initialization examples of GAC are illustrated in Fig. 7. We do admit that it is a drawback for a GAC to be initialized manually near the cell's contour edges. In fact, the GAC can only perform a semi-automatic contour extraction work, since the initial GAC must be given by an experienced/expert operator. But until now we have not found automatic GAC research work which can automatically set the initial GAC.

3.3. Comparisons

For the recognition comparisons, another scheme based on the gray and shape information is also performed [8]. The Freeman code is used in this scheme to extract the cell's contour edges. Then cell features including the cell's area, perimeter, centroid, profile and number are extracted by the image processing algorithms for the subsequent wood cell recognition. To ensure the objectivity/justice of experimental comparisons, the same image preprocessing procedure is used (i.e., for every wood cell species, 15 images are selected as the training cell images. Then the pattern recognition experiment is performed to the remaining 15 images. We also segment the original cell image into many small sub-images with a resolution 128×128). The relevant RRAs for five wood species *A. nephrolepis*, *Picea koraiensis*, *Pinus koraiensis*, *C. lanceolata*, *P. jezoensis* are approximately 80%, 75%, 85%, 77% and 88%, respectively, which indicates our scheme outperforms Ren's recognition algorithm in terms of the RRA.

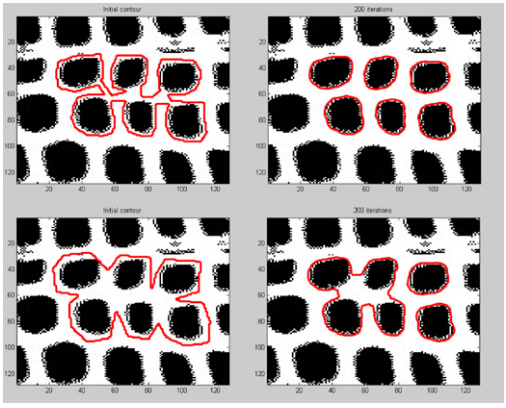

Fig. 7. Contour extraction comparisons by GAC with different initializations. (a) Good initial GAC; (b) good result after 200 iterations; (c) poor initial GAC; (d) poor result after 200 iterations.

4. Conclusions

In this paper, we propose a novel cell's shape recognition scheme that considers the wood cell's intraspecific and interspecific shape variations. This recognition scheme consists of two advantages. First, it can efficiently discriminate the intraspecific cell's shape variation and the interspecific cell's shape variation by using the PCA algorithm. Second, a GAC model is applied in a binary cell image, which can extract the edge contours of multiple cells simultaneously in a parallel way. Therefore, this scheme is not only time efficient by using a GAC model but also improves the cell recognition accuracy by using the PCA algorithm which can both effectively tolerate the intraspecific cell's shape variation and effectively identify the interspecific cell's shape variation.

Acknowledgments

This research is supported by the 2011 2011 National Natural Science Foundation with grant no. 31170515. It is also supported by the 2010 Heilongjiang Province Natural Science Foundation with grant no. F201005, the 2010 Heilongjiang Province Postdoctoral Foundation, and the Chinese Postdoctoral Science Foundation with grant no. 2011M500629.

References

[1] W. Gindl, H.S. Gupta, T. Schoberl, Mechanical properties of spruce wood cell walls by nanoindentation, Appl. Phys. A – Mater. Sci. Process. 79 (8) (2004) 2069–2073.
[2] B.D. Jordan, A simple image analysis procedure for fibre wall thickness, JPPS 14 (2) (1998) 44–49.
[3] S.P. Matti, S. Ritva, S. Andersson, Structural variation of tracheids in Norway spruce, J. Struct. Biol. 136 (2001) 101–109.
[4] M. Millin, W. Charles, Application of automatic image analysis to wood science, Wood Sci. 14 (3) (1982) 97–105.
[5] X. Diao, T. Furuno, T. Uehara, Analysis of cell arrangement in soft wood susing 2D fast Fourier transformation, Mokuzishi 42 (7) (1996) 634–641.
[6] T. Ona, T. Sonoda, In situ determination of proportion of cell types in wood by Fourier transform Raman spectroscope, Anal. Biochem. 268 (1999) 43–48.
[7] M. Yoko, M. Fujita, Experimental conditions for quantitative image analysis of wood cell structure-general procedures of Fourier transform image analysis, Mokuzai Gakkaishi 50 (2) (2004) 73–82.
[8] R.H.E., X.H. Tao, The characteristic parameter extraction of cell in board assortment recognition technology, J. Forest. Sci. 43 (9) (2007) 68–73 (in Chinese).
[9] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vision 1 (1988) 321–331.
[10] J.A. Sethian, Level Set Methods and Fast Matching Methods, Cambridge University Press, Cambridge, 1999.
[11] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vision 22 (1) (1997) 61–79.
[12] L.C. Ming, X.C. Yang, G.C. Feng, et al., Level set evolution without re-initialization: a new variational formulation, in: Proceedings of 2005 IEEE CVPR, vol. 1, 2005, pp. 430–436.