Computers and Electronics in Agriculture 109 (2014) 124–133

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

A novel matching algorithm for splitting touching rice kernels based on contour curvature analysis

P. Lin [a], Y.M. Chen [a], Y. He [b,*], G.W. Hu [a]

[a] College of Electrical Engineering, Yancheng Institute of Technology, No. 1 Middle Road Hope Avenue, Yancheng, Jiangsu Province 224051, PR China
[b] College of Biosystems Engineering and Food Science, Zhejiang University, 866 Yuhangtang Road, Hangzhou 310058, PR China

ARTICLE INFO

Article history:
Received 30 April 2014
Received in revised form 19 September 2014
Accepted 22 September 2014

Keywords:
Segment
Rice kernel
Contour curvature
Node matching

ABSTRACT

A novel node matching algorithm based on contour shape characteristics is introduced for accurately separating touching rice kernels. The original images of touching rice kernels are obtained by a scanner and preprocessed by de-noise, segmentation and contour extraction operations. The extracted contours are then smoothed by convoluting with Gaussian kernel function. Curvature analysis is used to detect characteristic touching points on the boundaries. The first-derivative of the curvature curve is taken to find its local peaks. The computed extremum of curvature correspond to the touching nodes along the original boundaries. Finally, the node matching rules including the confidence radiation region, the shortest distance, the length limitation of splitting line, etc., are proposed to determine an appropriate splitting line between related two of those nodes. The rules are key procedures for dealing with the problems of splitting complex touching kernels, and thus the process of how to determine the splitting line between touching kernels is detailedly discussed. One hundred scanning images with different shapes and sizes of rice kernels are used to estimate the robustness of the algorithm. Experimental results are encouraging that the proposed algorithm is not influenced by the exogenous parameters of rice kernels and can be used to effectively split kernels touching in a very complex way. The proposed methods can eliminate the traditional limitations of the manual placement of rice samples in a non touching manner before image acquisition and implement automatic system for the subsequent inspection of the appearance quality parameters of rice.

1. Introduction

Machine vision techniques have become popular in the investigation of rice quality in recent years (Costa et al., 2011; Shao et al., 2012; Liu & He, 2009; Wu et al., 2011). Numbers of algorithms of digital image process have been introduced to analyze the quality of rice, for example, the determination of the geometrical parameters (length, perimeter, projected area, etc), the broken ratio (Lloyd et al., 2001; Yadav and Jindal, 2001; Lin et al., 2012a,b), the whiteness (Liu et al., 1998a,b) and the fissures of rice kernel (Lan et al., 2002; Lin et al., 2012a,b). All previous works show perfect performance, however, one weakness of these studies was that the algorithms required the non-touching placement of all kernels. Under the touching condition, it is obvious that most of the former algorithms of feature extraction of kernels, i.e. the estimation of perimeter, broken ratio, etc. are prone to making mistakes. Therefore, before image collection, the grain kernels must be manually placed in a non-touching form. The procedures were quite inconvenient and time consuming. So it is necessary and urgent to develop an effective algorithm for splitting touching grain kernels in images, and thus the limitations of the placement of rice kernels can be eliminated in the process of image acquisition.

Significant progress has been made on the problems of segmentation of touching kernels in images. Shatadal et al. (1995) developed a segmentation algorithm based on image transform discipline of mathematical morphology. They tried to separate touching grain kernels through the erosion and dilation operations. The algorithm failed when a relatively long and narrow connected region were formed by the closely touching kernels. Visen et al. (2001) proposed a different approach based on the contour curvature analysis to segment occluding groups of grain kernels. The contour curvature values below a certain threshold were selected as splitting nodal points. The splitting lines between two of the nodes were determined only by the 'nearest-neighbor criterion', which limited to deal with the simple touching cases with maximum of three touching kernels. Zhang et al. (2005) employed an ellipse-fitting algorithm to separate touching grain kernels. Sample

* Corresponding author. Tel./fax: +86 571 88982143.
E-mail address: yhe@zju.edu.cn (Y. He).

http://dx.doi.org/10.1016/j.compag.2014.09.015
0168-1699/

points for fitted ellipses were found by randomly tracking the edge of touching kernels. The fitted ellipses were generated by a direct least-squares ellipse-fitting method, and then the clustering algorithm was employed to identify the best representative ellipse for each touching kernel. Finally, the touching grain kernels could be separated by the mathematical morphology transform. Such an algorithm demanded the kernels with approximately regular ellipsoid shapes, otherwise the splitting rate decreased. Wang and Chou (2004) developed an active contour model and inverse gradient vector flow (IGVF) field to segment touching rice kernels. IGVF was proposed to automatically generate a field center for individual rice kernels in an image. These centers were used as the references for setting initial deformable contours for building an active contour model. The algorithm could reconstruct the whole contour and achieve more than 96% similarity compared with the original contours. An alternative watershed transform approach has been reported for digital image segmentation. Initially, the splitting results were unsatisfying due to the over-segmentation reason (Bleau and Leon, 2000). Several approaches have been developed to solve the problems. However, it still had difficulties to separate the elongated grains in images (Casasent et al., 2002; Wang and Paliwal, 2006). In addition, those reported methods were also limited to the applications of splitting scenario of simple touching kernels. The watershed method was further improved by Zhong et al. (2009). Firstly, the distance and watershed transforms were used. Secondly, the watershed post-processing of over-segmentation was carried out by utilizing the convex features of related shapes. Thirdly, the candidates for splitting lines of touching clusters were found by matching the concavities and the rest of the un-segmented. Finally, several supplementary criterions were used to determine the splitting line. The results showed it could be employed to split the multiple clustered slender-particles. Mebatsion and Paliwal (2011) employed an elliptic Fourier series approximation to separate touching grain kernels. Nodal points were determined by the extrema of the contour curvature. A nearest-neighbour and radian critical (cumulative) distance difference of the chain-coded boundary point criteria was introduced to determine the segmentation lines between two nodal points. The algorithm appeared to be unsatisfactory when the kernels had rough boundaries or under the complicated touching condition. Park et al. (2013) introduced a robust approach that enabled automated morphology analysis of partially overlapping nanoparticles in electron micrographs. The splitting performance of the proposed algorithms showed pretty good. In Park's study, the contours of segmentation objects needed to be assumed to be convex, however, the contour of polished rice kernel nearby the germ has concave shapes. Thus, Park's algorithms were not suitable for the cases of segmentation of the rice kernels.

It is noteworthy that there were several splitting algorithms which were available for separation of rice kernels in images (Wang and Chou, 2004; Zhong et al., 2009), however, the current paper focuses on coping with the previously unsolved problems of splitting larger-scale touching kernels in images based on contour curvature analysis. To overcome the major splitting difficulties, novel matching rules including the confidence region of radiation, the shortest distance, the length limitation of splitting line, etc., for accurately determine the splitting line between two of the nodal points are proposed. As far as we know, this approach has not been reported, and that is the main motivation of writing this paper and the most significant parts of our work.

The paper is organized as follows. Firstly, experimental equipment and procedures are introduced. Secondly, methods of image preprocessing, smoothing, locating curvature peaks and node matching are presented. Thirdly, several illustrative examples of splitting simple and multiple touching rice kernels in images are discussed. Finally, draw conclusions.

2. Experiment and methods

2.1. Experiment

All of the experimental images were acquired by an HP LaserJet M1005 MFP with an optical scan resolution of 1200 dpi at 24 bits depth in grayscale style. The algorithms were performed in Matlab R2009a (The Math Works, Natick, USA) automatically.

Rice samples were placed on the glass plate of the scanner and covered with a piece of black board to acquire an image. Two kinds of touching images of rice kernels were investigated, including the simple and complex touching cases. Simple touching scenario was that just a few clusters were placed touched one by one and without holes (black backgrounds) inside shown in Fig. 1(a). Complex touching scenario was that dozens of rice kernels were placed closely, which led to some kernels and black holes surrounded by part of kernels shown in Fig. 1(b). One hundred images of slender and round rice kernels (each type contained 50 images) were collected for validation of the proposed algorithms. These images consisted of simple or complex touching cases. The details about the datasets of touching rice kernels were shown in Table 1.

2.2. Methods

2.2.1. Image preprocessing

Image segmentation is an important step in the image processing technique. The obtained scanning image is an intensity image. The intensity image is firstly converted into the binary image using the Otsu's method (Otsu, 1979; Moghaddam and Cheriet, 2012) which chooses the threshold to minimize the interclass variance of the black and white pixels. Image segmentation is an important step in the image processing technique. In our study, the contrast of intensity between the regions of the white rice samples and black background is strong, so the binarization segmentation errors might seldom arise. In spite of this case, the morphological smooth operation (e.g. a dilation following by an erosion using the same flat, disk-shaped structuring element with the radius of 8 pixels) has yet been implemented to remove the unforeseen noise from images. Due to implement of above strategy, the good performance of segmentation had been achieved. In practice, the segmentation errors might be mainly caused by the noise during imaging. It was suggested that the noise could be removed by using the classical filter technologies such as Gabor (Chen et al., 2011) and Wavelets functions (Lin et al., 2012a,b). It is inevitable that a little segmentation errors might appeared during the binarization process. These errors would add the noise on the contour curve. In order to remove the noise, the contour curves were

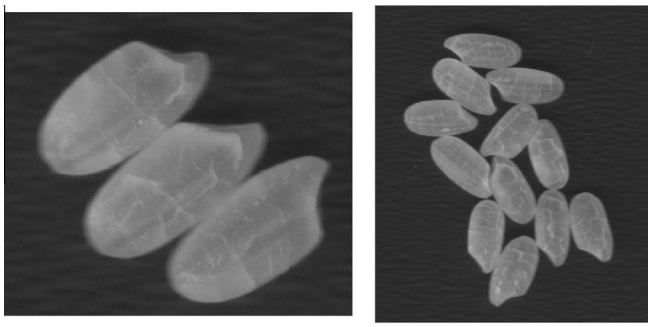

Fig. 1. (a) The simple touching scenario in which rice kernels touched one by one and (b) the complex touching scenario in which some rice kernels and black holes (black backgrounds) were surrounded by part of kernels.

Table 1
Details about the datasets of touching rice kernels.

Rice shape	Simple touching scenario		Complex touching scenario	
	Number of dataset	No. of clusters	Number of dataset	No. of clusters
Slender	2	1–5	5	20–50
	2	5–10	10	60–185
	2	10–15	8	200–380
	3	15–20	18	410–700
Round	2	1–5	7	20–70
	2	5–10	9	90–250
	2	10–15	11	280–490
	2	15–20	15	500–750

convoluted with a 1-D Gaussian kernel function. After that the segmentation errors would have little influence on the following splitting experiment. Finally, the boundaries function in Matlab R2009a is used to extract each region boundary of rice kernels in the binary image.

2.2.2. Discretely resampling the contours

There are two varieties of the contours, including the outside and inside contours. The outside refers to the outermost contour of touching rice kernels (see the red[1] circles in Figs. 7(a) and 8), and the inside refers to the contour of the black hole region enclosed by touching rice kernels (see the blue circles in Figs. 7(a) and 8). Assume that the extracted original contour curve is a single closed curve Γ. The curve Γ is discretely resampled and denoted as:

$$z(t) = (x(t), y(t)) \tag{1}$$

where x and y denote the abscissa and ordinate of points indexed by $t \in Z^+$ on curve Γ.

2.2.3. Smoothing the contours

The boundaries are represented by discrete points, thus there must be some noise in the contour curve Γ. The noises on the contour curve are removed by convoluting with a 1-D Gaussian kernel function $g(t,\sigma) = e^{-\left(\frac{t^2}{2\sigma^2}\right)}$:

$$\begin{aligned} Z(t) &= z(t) \otimes g(t,\sigma) = x(t) \otimes g(t,\sigma) + jy(t) \otimes g(t,\sigma) \\ &= X(t) + jY(t) \end{aligned} \tag{2}$$

where the symbol $\otimes$ denotes the convolution and $j^2 = -1$. σ is the kernel width.

2.2.4. Estimating curvature and curvature center

The curvature parameters are used to measure the convexity, concavity and bent level of the shape. The curvature of a curve is defined as the derivative of the slope of the tangent vectors expressed as (Mokhtarian and Abbasi, 2002):

$$\kappa(t) = (\dot{X}(t)\ddot{Y}(t) - \ddot{X}(t)\dot{Y}(t))/(\dot{X}^2(t) + \dot{Y}^2(t))^{3/2} \tag{3}$$

where $\dot{X}$, $\ddot{X}$ and $\dot{Y}$, $\ddot{Y}$ are the first and second derivatives of X and Y, respectively. The positive and negative curvatures denote that the contour shapes are convex or concave, respectively. The corresponding center of curvature $(X_C(t), Y_C(t))$ at the contour point $(X(t), Y(t))$ can be evaluated as following (Gu, 2007):

$$\begin{cases} X_C(t) = X(t) - \dot{Y}(t)(\dot{X}^2(t) + \dot{Y}^2(t))/(\ddot{Y}(t)\dot{X}(t) - \ddot{X}(t)\dot{Y}(t)) \\ Y_C(t) = Y(t) - \dot{X}(t)(\dot{X}^2(t) + \dot{Y}^2(t))/(\ddot{Y}(t)\dot{X}(t) - \ddot{X}(t)\dot{Y}(t)) \end{cases} \tag{4}$$

Thus, the convex and concave directions of a contour curve in an image can be determined by combing the contour point $(X(t), Y(t))$ with the corresponding curvature center of $(X_C(t), Y_C(t))$.

2.2.5. Finding curvature peaks

It is important to find the peaks of a curvature curve, because these points of peaks will be subsequently used to determine the splitting lines of touching rice kernels. The splitting points (SPs) can be found at the most convex or concave parts on the smoothed curvature curve $\kappa(t)$ and determined by taking the first derivative:

$$t_{MCCP} = \arg\{\dot{\kappa}(t) = 0 : |\kappa(t)| > \lambda\} \tag{5}$$

where $\lambda \in \Re^+$ is a threshold used to remove the noise in the curvature curve and determined by the sum of the mean value $E(\kappa)$ and the standard deviation of N curvature sample points:

$$\lambda = \begin{cases} E(\kappa) + \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(\kappa(t) - E(\kappa))^2}, & \text{for inside contour} \\ -E(\kappa) + \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(\kappa(t) - E(\kappa))^2}, & \text{for outside contour} \end{cases} \tag{6}$$

2.2.6. Locating MCCPs on the original contour

It is difficult to directly and accurately orientate the most convex or concave points (MCCPs) on the original curve, since there is too much noise on it and the derivative in Eqs. (3)–(5) is unable to be taken. One feasible way to determine the MCCPs on the original curve is to use the characteristics of the corresponding MCCPs on the smoothed curve. The estimated positions of the MCCPs on the smoothed curve do not correspond to the MCCPs on the original curve when using the same index t. The characteristic points on the original curve were drifted because the shape of the original curve was shrunk by the Gauss convolution function. So linking the splitting lines between such MCCPs will make mistakes. In order to accurately locate the MCCPs on the original curve and determine the splitting lines between touching kernels in the image, the following rules are proposed:

Firstly, build a new Cartesian coordinate system through the rotation and translation transformations as following:

$$\begin{cases} x' &= x\cos\phi + y\sin\phi - (x_0\cos\phi + y_0\sin\phi) \\ y' &= -x\sin\phi + y\cos\phi - (-x_0\sin\phi + y_0\cos\phi) \end{cases} \tag{7}$$

where $O' - x'y'$ is the transformed Cartesian coordinate obtained by shifting and rotating the original coordinate $O - xy$. (x_0, y_0) is the location of O' in the original coordinate. $\phi \in [0, 2\pi)$ is the rotating angle of the transformed coordinate along the counter-clockwise direction. Each MCCP on the smoothed contour is set as the origin of the transformed coordinate. The positive direction of y'-axis is from the MCCP to its curvature center D, and the corresponding positive direction of x'-axis is determined by rotating the y'-axis

[1] For interpretation of color in Figs. 6–8, the reader is referred to the web version of this article.

by an $\pi/2$ angle clockwisely. As shown in Fig. 2, the position of B on the original curve is distributed between two points of A and C along the original curve on the $x' \geqslant 0$ and $x' < 0$ half planes of the transformed coordinate, respectively. The indexes of points of A and C can be estimated by the following equations, respectively:

$$\begin{cases} t_A = \arg\min_{t\in\Delta}\{\|(x'(t),y'(t)) - (0,0)\|_2 : |y'(t)| < \lambda, x'(t) \geq 0\} \\ t_C = \arg\min_{t\in\Delta}\{\|(x'(t),y'(t)) - (0,0)\|_2 : |y'(t)| < \lambda, x'(t) < 0\} \end{cases} \quad (8)$$

where $\|\bullet\|_2$ is the L2-norm and λ is a threshold that used to limit the number of the estimated points near the x'-axis. Δ represents the point set between the points of A and C on the concave part. The point of B has the longest distance d to the x'-axis than other points in the set Δ, so it is considered as the MCCP. Corresponding index of point B can be evaluated by:

$$t_{MCCP} = \arg\max\{|y'(t)||t \in [\min(t_A, t_C), \max(t_A, t_C)]\} \quad (9)$$

The flow chart of exactly locating the MCCP on the original contour was shown in Fig. 3.

2.2.7. Matching characteristic points

After obtaining the MCCPs, the next step is to match these characteristic points to draw the splitting line in the image. The characteristic points are divided into two varieties, one is the base point (BP) such as the red dot O' and the other is the matching point (MP) such as red dots O'' and O''' (see Fig. 4). The BP was usually randomly selected, but the determination of the appropriate MP need to meet certain rules. Such rules are listed as follows:

- Firstly, the candidate for MPs must be in the triangular radiation region, which is covered with horizontal dash-line net in Fig. 4. The radiation region derives from the BP and its normal direction is opposite from the BP to its curvature center D, whose direction is the same as the vector $\overline{DO'}$ in Fig. 4. The radiation angle is 2θ, which is formed by rotating the extend line l_1 of $\overline{DO'}$ at angle θ clockwise and counter-clockwise along the point O', respectively. The triangular radiation region and angle 2θ are called the matching confidence triangular radiation region (CTRR) and angle (CTRA), respectively. The point of O''' in the CTRR is considered as a possible MP to the BP O'. The MP O'' out of the CTRR cannot be used as corresponding MP. Besides the CTRR, the CTRA can also be used to determine the MP. When the angle between the normal line and linking line between the MP and BP is less than half of the CTRA of 2θ, the corresponding MP is considered as a correct SP. As shown in Fig. 4, the angle α between line l_1 and l_2 is less than half of the CTRA, so the MP of O''' is in the CTRR and may become the correct SP. The angle β is larger than half of the CTRA, so the MP O'' is not correctly matched. The angle $\alpha \in [0, \pi)$ between line l_1 and l_2 in the transformed coordinate can be computed as follows:

$$\alpha = \arctan\left((-1)^{Heaviside} \cdot (k_2 - k_1)/(1 + k_1 k_2)\right) + \pi \cdot Heaviside \quad (10)$$

$$Heaviside = \begin{cases} 0, & x\cos\phi + y\sin\phi - (x_0\cos\phi + y_0\sin\phi) \geq 0 \\ 1, & x\cos\phi + y\sin\phi - (x_0\cos\phi + y_0\sin\phi) < 0 \end{cases} \quad (11)$$

where k_1 and k_2 are the slopes of the line l_1 and l_2, respectively. (x_0, y_0) and (x, y) are the coordinates of the BP and MP of O' and O''', respectively. ϕ is the angle between x'- and x-axis.

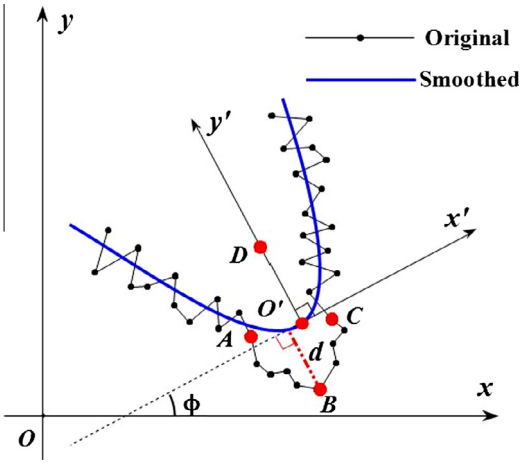

Fig. 2. Characteristic points on the original and smoothed boundary.

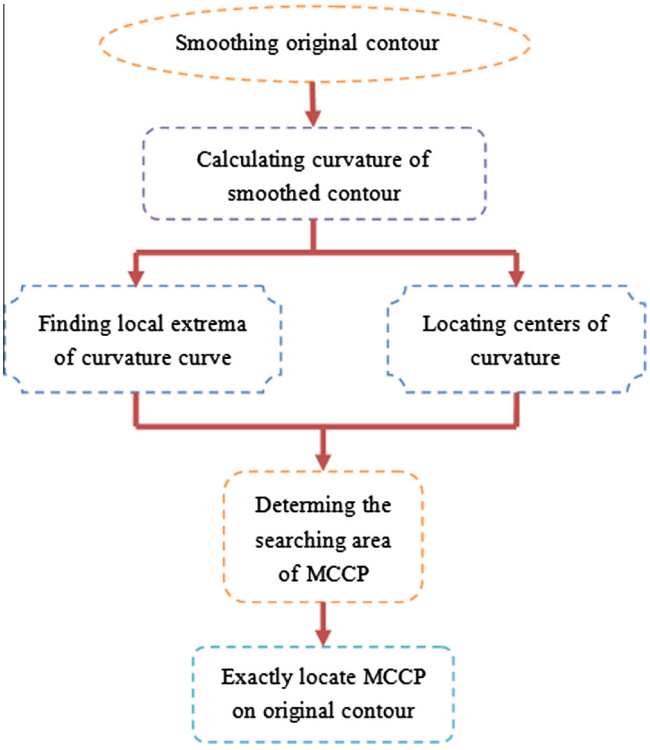

Fig. 3. Flow chart of locating the most convex and concave points (MCCPs) on the original contour.

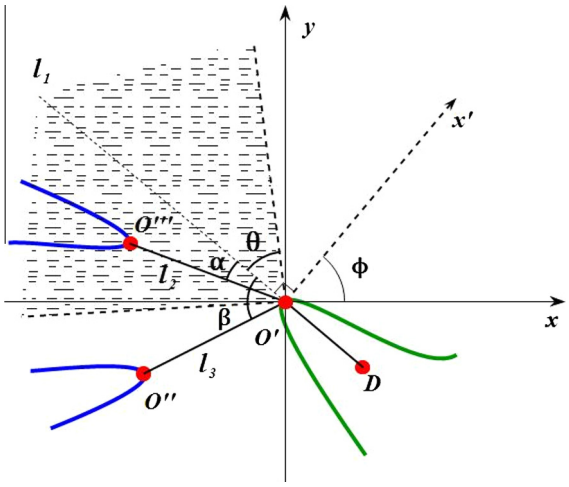

Fig. 4. Determining the matching points using the confidence triangular radiation region.

- Secondly, any two convex points extracted from the boundary of the same black background patch cannot be linked together (see red points on the blue contours marked in Figs. 7(a) and 8), since the formed splitting lines just separate the background but not touching kernels. However, any two convex points on the outermost boundary (see yellow points on the red contours marked in Figs. 7(a) and 8) or any two convex and concave points can be applied to matching.
- Thirdly, two points obtained from the previous screening procedures with the shortest distance is firstly considered as a pair of the correct SPs and can be determined by:

$$\{t_i, t_j\} = \arg\min_{\varsigma,\xi\in\Delta}\{\|(x(\varsigma), y(\varsigma)) - (x(\xi), y(\xi))\|_2 | \varsigma \neq \xi\} \quad (12)$$

where $\|\bullet\|_2$ denotes the 2-norm. Δ represents the point set. ξ and ς are the index of the BP and MP respectively. The locations of the SPs are $(x(t_i), y(t_i))$ and $(x(t_j), y(t_j))$ in the image.

- Fourthly, the length of the splitting line must be less than the longest length of intact rice. The limitation can prevent error matching to some extent and was computed as following:

$$d < C \cdot \max_{i\in[1,n]}\{L(i)\} \quad (13)$$

where d is the distance between two SPs. $L(i)$ is the maximum length between two pixel points on the contour of the i-th intact rice. n is the number of the rice kernels. $C \in \Re^+$ is a constant.

- Finally, after finishing the above first-run matching procedures (steps 1–4), there may still exist few unmatching points (the phenomenon will be shown in the following section of complex touching scenario). The rest of the points are set as a new BP and matched again by redoing steps 1–4.

The flow chart of matching the MCCPs is shown in Fig. 5.

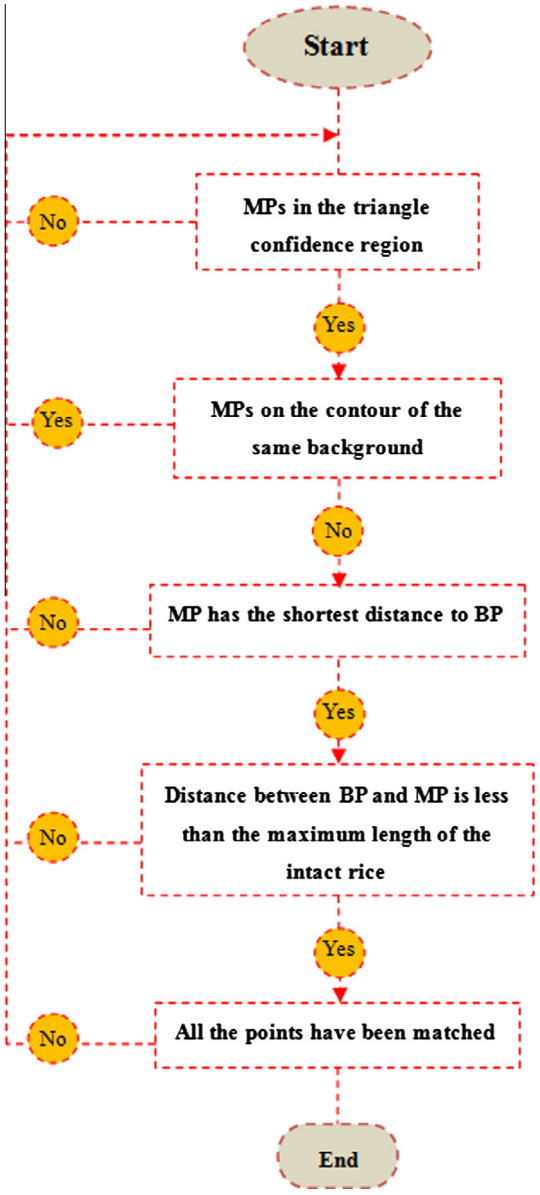

Fig. 5. Flow chart of matching the MCCPs.

3. Results and discussion

In this section, the matching methods were introduced for splitting two types of touching cases of rice kernels: one was the simple touching scenario and the other was the complex touching scenario which was the highlight of our research.

3.1. Simple touching scenario

The simple touching scenario of rice kernels was firstly used to test our methods for segmentation. As shown in Fig. 6(a), the black and red close curve denoted the original and smoothed contour, respectively. The red curve was obtained from the black curve by convoluting with the Gaussian kernel function with $\sigma = 30$. Eq. (3) was used to evaluate the curvature of the smoothed contour. There were four extrema obtained from taking the first derivative of the contour curvature by using Eq. (5). The extrema were marked by the red dots $(A - D)$ at the local peaks in Fig. 6(b) and corresponding concave points on the red smoothed contour were marked by the green dots $(A - D)$ in Fig. 6(a). The green dots $(A - D)$ were all located at the most concave positions on the smoothed contour. As shown in Fig. 6(a) and (b), the height of curvature peaks reflected the concave degree of the contour curve. For example, the contour curves around the dot B and C had the most and least concave shapes, which corresponded to the highest and lowest peaks on the curvature curve. When the same indexes of characteristic points $A - D$ on the smoothed contour are used to locate the points on the original curve, it was discovered that the corresponding characteristic points of $A' - D'$ drifted from the most concave points on the original contour (see Fig. 6(a) and (b). The drifting phenomenon was caused by changing the phase of the original contour during the Gaussian convoluting process. So the locations of local peaks of the curvature of the smoothed contour could not directly reflect the most concave positions on the original contour. To solve this problem, Visen et al. (2001) proposed an approach to segment occluding groups of grain kernels based on contour curvature analysis. They re-computed curvature along the boundary pixels using a 25% of frequency width. The bias threshold was determined empirically, so the value was easily alternated when using Gaussian kernel function with different widths, the scanners with different resolutions or by different persons, etc.

To exactly determine the drifting value in different situation, a novel matching method was introduced. Firstly, determine the seeking region for the most concave point on the original contour. As shown in Fig. 6(c), the position of the most concave point of A was between the points of A'_1 and A'_2, so the region between them could be used. The points of A'_1 and A'_2 were located by estimating the shortest distance from the nearest concave point A on the smoothed contour using Eq. (8), respectively. The point of A_0 had the longest distance to the tangent line (see the cyan dot-line in

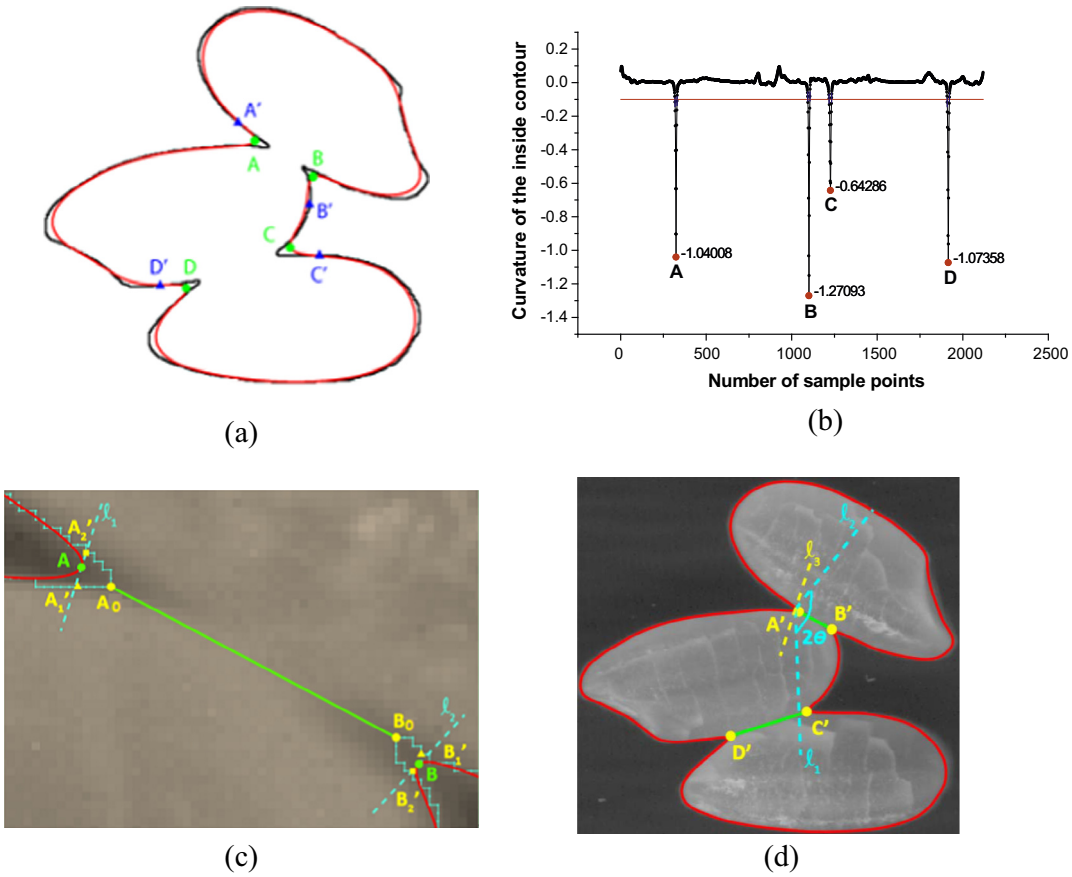

Fig. 6. (a) The concave points on smoothed boundary and the corresponding drifting points on the original boundary. (b) Curvature of the smoothed contour. (c) Exactly locating the concave points on the original contour. (d) Separating touching rice kernels.

Fig. 6(c)) that through the point A, which was considered as the most concave point on the original contour. Eq. (9) was used to compute the index of A_0, and then the location of the point A_0 in the image was determined. Other most concave points such as $B' - D'$ on the original contour could be found by using the same methods (see Fig. 6(d)). After finding the most concave points on the original contour, the touching rice kernel could be separated by linking two of the matching points. As shown in Fig. 6(d), when the point A' was selected as the BP, the rest of the three points of $B' - D'$ became the MPs. In order to find the correct MP for the BP, a CTRR deriving from the BP of A' with a large CTRA was suggested to be used to hunt for the MPs. As shown in Fig. 4, the direction of the CTRA was determined by rotating the tangent line by $\pi/2$ angle along the BP. The position of tangent point changed in consequence of the contour deformation during the Gaussian smooth procedure, which caused the direction drift of the CTRA to some extent. Thereby, the CTRA should not be set up too small in practice otherwise the MPs might go beyond the CTRR. The CTRA was supposed to be larger than $\pi/2$ angle. On the other hand, although the smooth process caused the drift of the tangent line, the direction of tangent line of the MP varied slightly due to the small change of the contour shape. Generally, the drift angle of the tangent line was less than $\pi/6$ angle was used to eliminate the influence of the drift. On the basis of these two points, it was suggested that the fitting value of the CTRA should be generated in the region of $[\pi/2, 5\pi/6]$. In this paper, the value of the CTRA was taken by $2\theta = 7\pi/9$. Obviously, the point D' was not in the CTRR, so it was firstly excluded. The points of B' and C' were in the CTRR, so they became the matching candidates. In terms of the proposed algorithms in the step three, the MP with the shortest distance to the BP would be finally used to determine the splitting line. The distance between the points of A' and B' was shorter than A' and C', so the point C' was excluded. The touching rice kernels could be separated by linking the splitting line between the points of A' and B'. The same procedures were repeatedly carried out until the remains of points were all exactly matched.

3.2. Complex touching scenario

Mebatsion and Paliwal (2011) showed that their matching methods could be used to deal with the complex cases of splitting more than three touching kernels, however, the placement patterns of the multiple kernels were specific in their experiments, where the number of the touching kernels was no more than three. As far as we knew, the algorithms based on the curvature analysis for dealing with the complex touching situation had not been implemented yet (Shatadal et al., 1995; Visen et al., 2001; Zhang et al., 2005; Mebatsion and Paliwal, 2011), however, they were needed in the practical applications and believed to be much more significant. In the following sections, our matching algorithms would be introduced to solve such complex cases.

Firstly, all the original outside and inside contour curves were smoothed by convolution with Gaussian kernels using Eq. (2). The perimeters of the outside and inside contours were different, so the filtering kernel widths σ were distinct. The smoothed contour curves could be used to compute the curvature. The smoothed outside and inside contours (holes) were marked by the red and blue lines, respectively (see Fig. 7(a)). When the same filtering width were used for the original curves it will lead to serious distortion of the contour curves of several small background blocks,

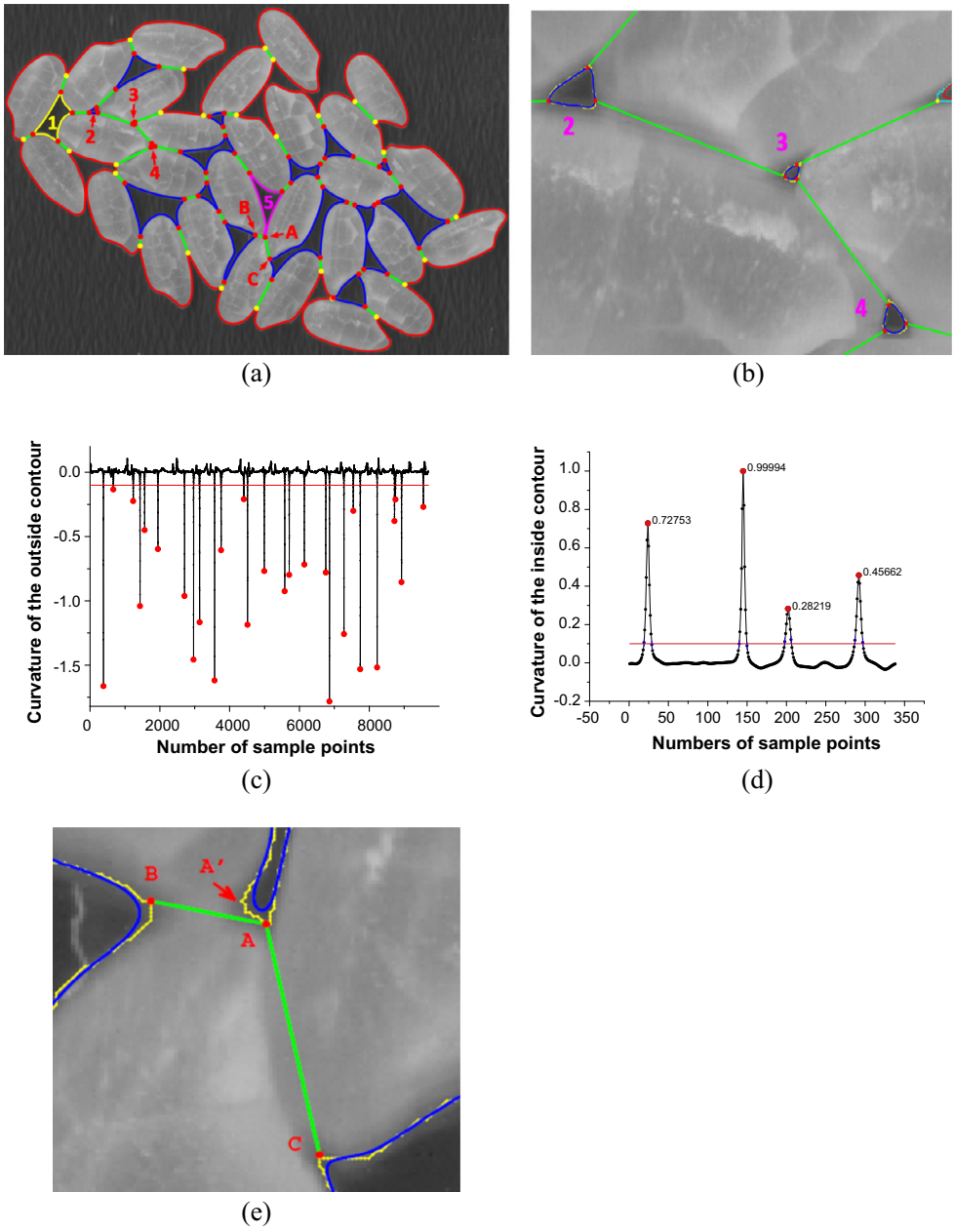

Fig. 7. (a) Separate multiple touching rice kernels with black background holes inside. (b) Smoothing small inside holes with the relative narrow Gaussian kernel width. The curvature curves of the outside (c) and inside contours (d) and the peaks corresponding to the MCCPs on the smoothed contours are marked with red dots. (e) Point A was matched twice with points B and C, for the characteristic point A′ being filtered out during Gaussian kernel convolution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

i.e. the curve became completely convex without any characteristic points on it. So the smoothed parameter σ should be tuned when dealing with the different perimeters of curves:

$$\sigma = \xi / \log_2(L) \tag{14}$$

where L is the length of the discrete boundary and ξ is a constant coefficient. In this study, L was divided into four groups of $[1, 80]$, $(80, 200]$, $(200, 2000]$ and $(2000, +\infty)$, and the corresponding coefficient ξ was set as 10, 30, 50 and 65. By using the width values of Eq. (14), even small characteristic points on the background contours could be detected (shown in Fig. 7(b)).

After obtaining the smoothed curves, the curvature of contours could be evaluated. Two different types of the contour curvatures (see Fig. 7(c) and (d)) were discussed here, one referred to the outside contour marked with the red circle, and the other referred to the one of the inside ones marked with the yellow circle. The Eq. (5) with the threshold value of $\lambda = 0.1$ was used to find the local peaks of the curvature curves. The absolute values of the local peaks of the curvature curves which were larger than the λ will be considered. Twenty-seven local minima and four local maxima points were found in the curvature curves and marked with the red dots in Fig. 7(c) and (d), which implied there were twenty-seven and four MCCPs on the outside and selected inside smoothed contours, respectively. The corresponding locations of the computed MCCPs were marked by the yellow and red dots in Fig. 7(a) and 8, respectively.

To determine the splitting lines between the touching kernels, it is necessary to determine the properties of the node positions at

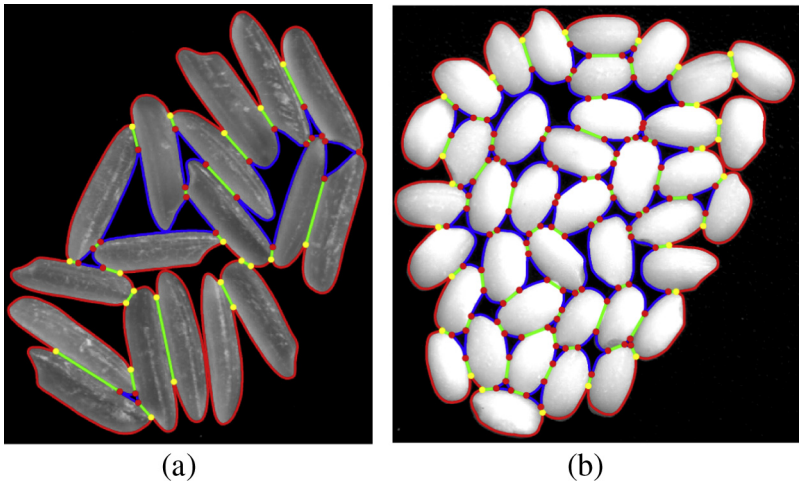

Fig. 8. Splitting the clustered (a) slender and (b) round rice kernels.

first. As shown in Fig. 7(a), the nodal points were divided into two types: the outside and inside according to their positions on the outermost and inside contours, respectively. Thereby, the splitting line might be connected between the outside and inside, inside and inside, and outside and outside nodal points. As shown in Fig. 7(e), the inside dot *A* was matched twice: one is with the dot *B*, and the other is with the dot *C*. The dot *A* was matched with the dot *C* at the first matching process. The dot B was unmatched for outside the CTRR from the BP *A*. According to the matching rules in the step five, the remainder of the unmatched points could be used as the BP at the next matching procedure. When the point *B* was used as the BP, the point *A* was found in the CTRR and had the shortest distance to *B*, so the points *A* and *B* became a pair. So the point *A* was linked twice shown in Fig. 7(e). However, the point *A*′ rather than *A* was true matching dot of point *B*. The deviation occurred because the number of the discrete sampled points on the inner

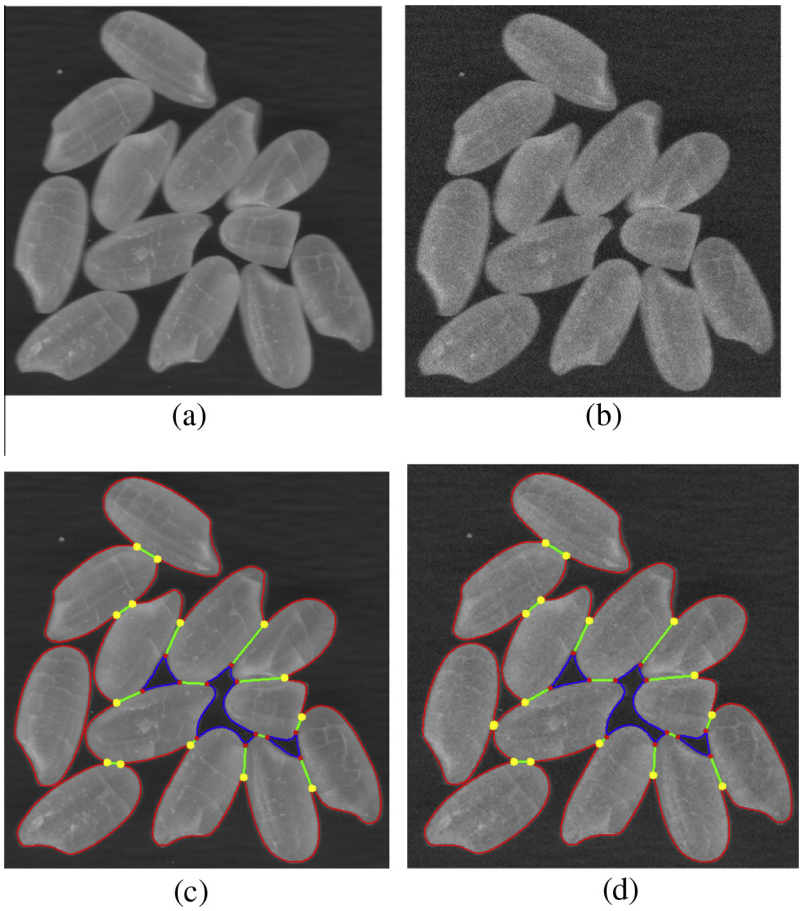

Fig. 9. Comparison of segmentation performance of algorithms in ideal and noise contaminated environment. (a) Origin image. (b) Speckle noise contaminated image, PSNR = 25 dB and SSIM = 0.325. (c) Splitting the ideal image. (d) Splitting the contaminated image.

contour is small due to the low resolution of the equipment, and the characteristic point A' was inevitably removed during the smoothed process. The distance between the points A and A' was extremely short, so the splitting lines of BA and BA' could be considered as approximately coinciding with each other to a certain extent, and the linking line between the points of A and B could be considered as the correct splitting line.

Two different shapes of rice kernels with different sizes were employed to test the robustness of the splitting algorithm. The results of the segmentation of the clustered slender and approximately round rice kernels were shown in Fig. 8(a) and (b), respectively. The experimental analysis of the quality of segmentation of processed images was determined by:

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (15)$$

where TP and FP is the true and false positives referring to the number of the true and false separating lines, respectively. For all the 100 images, Thirty-seven nearly coincide points (see Fig. 7(e)) were detected, and the splitting lines linked were considered as correct splitting points (see Section 3.2), Other connected cases were all correctly splitted. The separating results showed the precision rate reached 100%. It was encouraging that the proposed approaches were not affected by the exogenous parameters and could be used to implement accurate segmentation of the touching kernels across whole dataset even under extremely complex touching environment. The computational efficiency of our method was demonstrated by Monte-Carlo experiment. The simulation results showed that the average separating time for 100 touching rice kernels was about 4.3 s when 1000 Monte-Carlo tests were carried out. The configulation of the computing circumstance was Windows 8 Intel(R) Core(TM) i5-4200U CPU @ 1.60 GHz DDR3L 1600 MHz 4G.

3.3. Splitting the contaminated images

The robustness of the algorithms was evaluated when the quality of image degraded moderately. The cases were simulated by adding the Gaussian and speckle noises to the original scanning images. The contaminated level was quantified using peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) index. The typical values for lossy compression of an image were between 30 and 50 dB and when the PSNR was less than 40 dB the two images are distinguishable (Kawahito et al., 1997). Wang et al. (2004) showed that the perceptual quality of image drastically degraded when the SSIM index was less than 0.7. In this study, the thresholds of PSNR and SSIM for each image were adjusted to 30 dB and 0.7. The symlet wavelet with the Donoho–Johnstone universal threshold was performed to smooth images before segmentation (Adamo et al., 2013). As shown in Fig. 9, the contaminated images blur the adjacent regions of approximately connected rice kernels. Such regions might become connected after the filtering process. However, the results demonstrated that the performance of segmentation algorithms was not affected by the low quality of images (see Fig. 9(d) at the location of A). All the images from the database were used for test. The precision rate also reached 100%, which demonstrated the robustness of the segmentation algorithms.

4. Conclusions

The proposed approach overcomes the restrictions of the existing methods like demanding the kernels with approximately regular ellipsoid or convex shapes (Park et al., 2013), limitation to deal with the simple touching cases with maximum of three touching kernels (Mebatsion and Paliwal, 2011; Visen et al., 2001), or the complexity of the touching cases (Shatadal et al., 1995; Mebatsion and Paliwal, 2011). The paper introduced a novel matching algorithm for improving the performance of splitting touching rice kernels of different shapes and sizes in images based on the contour curvature analysis. The experiment results illustrated that the proposed methods could be used to split the simple touching scenarios of rice kernels, but such a simple situation was far from the practical application of grain checking. This research went further into the case of separating complicated scenarios of the touching kernels based on the curvature analysis, which had not been reported in the previous related investigations. In order to validate the robustness of the splitting algorithm, one hundred scanning images of rice kernels with different shapes and sizes were used for test. The experiment results were encouraging that the splitting algorithms were not influenced by the exogenous parameters. Therefore, the proposed methods could be used as the pretreatment of the subsequent rice quality inspection. Combining with such preprocess, the synthesized algorithm had the potential to enhance the detecting accuracy and efficiency of the rice quality by using machine vision. Moreover, there were numerous possible extended applications of our methods to split touching instances of other grain kernels such as soybean, wheat and corn. There was a need to point out that although our methods could be successfully used to segment touching kernels in images, the patterns of placement of the rice kernels still had certain restrictions, i.e. all kernels were required to be placed in a single layer during imaging. Therefore, in our future research, we will consider the larger-scale multi-level touching conditions and provide a more practical and powerful algorithm for segmentation.

Acknowledgments

This study was supported by Natural Science Research Project of Higher Education of Jiangsu Province (Grants No. 13KJB210006) and the Natural Science Foundation of Jiangsu Province (Grants No. BK20140467). Yancheng Institute of Technology Breeding Programs (Grants No. KJC2014006, KJC2014007, XKY2014055, XKY2014056).

References

Adamo, F., Andria, G., Attivissimo, F., Lanzolla, A.M.L., Spadavecchia, M., 2013. A comparative study on mother wavelet selection in ultrasound image denoising. Measurement 46 (8), 2447–2456.
Bleau, A., Leon, L.J., 2000. Watershed-based segmentation and region merging. Comput. Vis. Image Underst. 77 (3), 317–370.
Casasent, D., Talukder, A., Keagy, P., Schatzki, T., 2002. Detection and segmentation of items in X-ray imagery. Trans. ASAE 44 (2), 337–345.
Chen, Y.M., Lin, P., He, Y., Xu, Z.H., 2011. Classification of broadleaf weed images using Gabor wavelets and lie group structure of region covariance on riemannian manifolds. Biosyst. Eng. 109 (3), 220–227.
Costa, C., Antonucci, F., Pallottino, F., Aguzzi, J., Sun, D.W., Menesatti, P., 2011. Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision. Food Bioprocess Technol. 4 (5), 673–692.
Gu, J.M., 2007. Conic's center of curvature's trochoid equation. J. Guangsha College Appl. Constr. Technol. 1, 1–3 (in Chinese).
Kawahito, S., Yoshida, M., Sasaki, M., Umehara, K., Miyazaki, D., Tadokoro, Y., Matsuzawa, A., 1997. A CMOS image sensor with analog two-dimensional DCT-based compression circuits for one-chip cameras. IEEE J. Solid-State Circuits 32 (12), 2030–2041.
Lan, Y., Fang, Q., Kocher, M.F., Hanna, M.A., 2002. Detection of fissures in rice grains using imaging enhancement. Int. J. Food Prop. 5 (1), 205–215.
Lin, P., Chen, Y.M., Bao, Y.D., He, Y., 2012a. Image detection of rice fissures using biorthogonal B-spline wavelets in multi-resolution spaces. Food Bioprocess Technol. 5 (5), 2017–2024.
Lin, P., Chen, Y., He, Y., 2012b. Identification of broken rice kernels using image analysis techniques combined with velocity representation method. Food Bioprocess Technol. 5 (2), 796–802.
Liu, F., He, Y., 2009. Application of successive projections algorithm for variable selection to determine organic acids of plum vinegar. Food Chem. 115 (4), 1430–1436.

Liu, W., Tao, Y., Siebenmorgen, T.J., Chen, H., 1998a. Digital image analysis method for rapid measurement of rice degree of milling. Cereal Chem. 75 (3), 380–385.
Liu, W., Tao, Y., Siebenmorgen, T.J., Chen, H., 1998b. Digital imaging method for rapid measurement of rice degree of milling. Rice Res. Ser. 460, 111–116.
Lloyd, B.J., Cnossen, A.G., Siebenmorgen, T.J., 2001. Evaluation of two methods for separating head rice from brokens for head rice yield determination. Appl. Eng. Agric. 17 (5), 643–648.
Mebatsion, H.K., Paliwal, J., 2011. A Fourier analysis based algorithm to separate touching kernels in digital images. Biosyst. Eng. 108 (1), 66–74.
Moghaddam, R.F., Cheriet, M., 2012. AdOtsu: an adaptive and parameterless generalization of Otsu's method for document image binarization. Pattern Recogn. 45 (6), 2419–2431.
Mokhtarian, F., Abbasi, S., 2002. Shape similarity retrieval under affine transforms. Pattern Recogn. 35 (1), 31–41.
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst., Man, Cybernet. 9 (1), 62–66.
Park, C., Huang, J.Z., Ji, J.X., Ding, Y., 2013. Segmentation, inference, and classification of partially overlapping nanoparticles. IEEE Trans. Pattern Anal. Mach. Intell. 35 (3), 669–681.
Shao, Y.N., Zhao, C., Bao, Y.D., He, Y., 2012. Quantification of nitrogen status in rice by least squares support vector machines and reflectance spectroscopy. Food Bioprocess Technol. 5 (1), 100–107.
Shatadal, P., Jayas, D.S., Bulley, N.R., 1995. Digital image analysis for software separation and classification of touching grains: I disconnect algorithm. Trans. ASAE 38 (2), 635–643.
Visen, N.S., Shashidhar, N.S., Paliwal, J., Jayas, D.S., 2001. AE-automation and emerging technologies. J. Agric. Eng. Res. 79 (2), 159–166.
Wang, Y.C., Chou, J.J., 2004. Automatic segmentation of touching rice kernels with an active contour model. Trans. ASAE 47 (5), 1803–1811.
Wang, W., Paliwal, J., 2006. Separation and identification of touching kernels and dockage components in digital images 48 (7), 1–7.
Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612.
Wu, D., Nie, P.C., He, Y., Bao, Y.D., 2011. Determination of calcium content in powdered milk using near and mid-infrared spectroscopy with variable selection and chemometrics. Food Bioprocess Technol. 5 (4), 1402–1410.
Yadav, B., Jindal, V., 2001. Monitoring milling quality of rice by image analysis. Comput. Electron. Agric. 33 (1), 19–33.
Zhang, G., Jayas, D.S., White, N.D.G., 2005. Separation of touching grain kernels in an image by ellipse fitting algorithm. Biosyst. Eng. 92 (2), 135–142.
Zhong, Q., Zhou, P., Yao, Q., Mao, K., 2009. A novel segmentation algorithm for clustered slender-particles. Comput. Electron. Agric. 69 (2), 118–127.