

ESTIMATION OF LOSS RATE OF OATS CLEANING BASED ON WATERSHED SEGMENTATION

基于分水岭分割的燕麦清选损失率估计研究

Hongwen Yan, Qingliang Cui¹, Xuefeng Deng¹

College of Information Science and Engineering, Shanxi Agricultural University, Taigu/China

Tel: +86-0354-6289253; E-mail: qlcui@126.com

DOI: 10.35633/INMATEH-59-14

Keywords: oats, cleaning, watershed segmentation, loss rate

ABSTRACT

This paper studied the loss rate of oats in the process of cleaning from the perspective of image processing. The sample was divided into group a that contained no impurities and group b that contained impurities. Otsu method was used to segment the oat kernels, with the recognition rate reaching 94.20%, and morphological opening was used for the openings appearing during the segmentation process for filling, while watershed segmentation algorithm was used for segmentation of adhesion area, with the recognition rate reaching 98.50%. For group b, the area method was used to identify and separate the impurities. Through statistical analysis, the area threshold was 600 pixels, and impurities could be removed without excessive segmentation. The estimated 5 g-sample loss rate in group a was 2.08%, which met requirements, so 5 g-sample was selected in group b, and it was calculated that the estimated loss rate of group b was 2.60%. The study showed that having good effect on image processing with less adhesion after cleaning, the algorithm could provide theoretical and methodological support for on-line monitoring of loss rate during oats cleaning.

摘要

本文从图像处理角度研究燕麦在清选过程中的损失率。将样品分为不含杂质的 a 组和含杂质的 b 组，使用 Otsu 方法分割燕麦籽粒，识别率达 94.20%。并对分割过程中的孔洞使用形态学开运算填充，对粘连区域采用分水岭分割算法进行分割，识别率达 98.50%。对 b 组采用面积法对杂质识别并分离，经统计分析，面积阈值为 600 个像素，可去除杂质且不会出现过度分割。a 组中 5g 样品损失率估计值为 2.08%，满足要求，故在 b 组选取 5g 的样品，计算 b 组损失率估计值为 2.6%。研究表明，该算法对于清选后粘连较少的图像处理效果好，可对燕麦清选损失率的在线监测提供理论和方法支持。

INTRODUCTION

Oats are traditional health food that has an irreplaceable role in stabilizing food production and maintaining dietary diversity in drylands. The loss of grain is unavoidable in the process of cleaning during oat harvesting (Tang et al., 2013; Wei et al., 2016). The loss rate is an important technical indicator for measuring the performance of the cleaning machine and an important basis for the adjustment of the working parameters of the cleaning machine (Liang et al., 2015; Zhang et al., 2012). Most of the calculation of the loss rate was manually calculated (Zhou et al., 2010), which was low in accuracy and inefficient. It is a prerequisite to study the algorithm for automatic calculation of the loss rate during the cleaning process for the development of intelligent cleaning equipment.

The use of machine vision on the quality of kernel of bulk grain had been studied in depth at home and abroad (Diego and Rafael, 2018; Uryi et al., 2018; Vithu and Moses, 2016), but the researches on loss rate were few and most of the previous researches focused on the use of sensors (Eric, 2018; Gao et al., 2011; Mao et al., 2012). Some studies have focused on monitoring grain harvest losses through temperature differences captured by thermal sensor arrays, and optimizing sensor size and layout to improve signal recognition accuracy for rice loss monitoring (Liang et al., 2017), and designing a real-time monitoring system for losses during the cleaning process (Zhou et al., 2010), which could effectively differentiate the signals of

¹ Hongwen Yan, As. Ph.D. Eng.; Qingliang Cui, Prof. Ph.D. Eng.; Xuefeng Deng, Ph.D. Eng.

clods, stalks and stones in wheat, however, there were few reports on the calculation of the loss rate by directly obtaining the amount of oat kernels through images.

To explore the loss rate during the oat cleaning process, in this study, the cleaning machine developed by Shanxi Agricultural University was used as the test bench, oats of different weights were used as experimental samples, and oats going through the cleaning process of the test bench were collected. Watershed segmentation algorithm was adopted to count the number of oats after cleaning, via which their weight could be estimated to offer support for online monitoring of loss rate of oats during cleaning.

MATERIALS AND METHODS

Sample collection

The experimental materials of this study were collected in Youyu County, Shanxi Province, China (112°27' E, 39°59' N), and the sampling date was July 10, 2018. Since the cleaning machine would inevitably have some impurities after the cleaning operation was completed, in order to facilitate the comparison of effects of the algorithm, the sample was divided into group a that contained no impurities and group b that contained impurities. Impurities referred to oat stalks, twigs, hulls, and oat leaves in the sample, as shown in Fig.1.



Fig. 1 - Images for Impurities in oat

Equipment

In the study, the cleaning machine developed by Shanxi Agricultural University was used as the test bench, and the oat samples taken were collected via this test bench. The speed used in the test was 68 rpm.

The image acquisition system used M0814-MP type lens (Computar, $f = 8\text{mm}$, F1.4), F201C type CCD (Allied vision technologies Gmb H, 1/2.3 inch CCD), RL500W type circular diffuse reflection LED light source (by Victor Digital Image Technology Co., Ltd. with polarizer, 9W), CP502 electronic balance (by Ohaus, with sense quality of 0.01 g) and R2004 tripod (by Si Rui Group, with a working height of 48.5~1 740 mm) as equipment.

Image acquisition

The oat sample fell from the screen box of the cleaning machine to the buffer receiving board on the floor. In order to reduce the shadow of the cleaning machine when the image was acquired, the buffer board was carefully removed from the lower end of the cleaning machine in this test during preparation for image acquisition.

- (1) Mount the light source and CCD camera on a tripod and fix them;
- (2) Adjust the electronic balance to zero in advance and clean the weighing environment;
- (3) Make calibration when recording image data to ensure that the image data acquired are uniform.

Data analysis software

The data processing and analysis in this study was based on the software Matlab 7.5 (The MathWorks, Natick, USA).

STATISTICAL ANALYSIS OF THE NUMBER OF OAT KERNELS

The sample collected in this study was divided into oat kernel sample that contained no impurities and oat kernel sample that contained impurities. The basic process of the study is as shown in Fig.2. The acquired image was preprocessed; the median filtering was used to reduce noise; the image was pre-segmented by Otsu method; the holes in the binarized image were filled; the watershed segmentation algorithm was used to segment the adhered oats, and oat image containing impurities was identified and separated by area method.

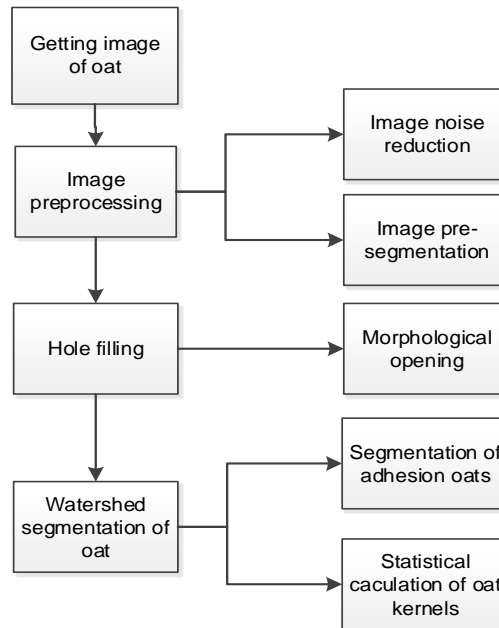
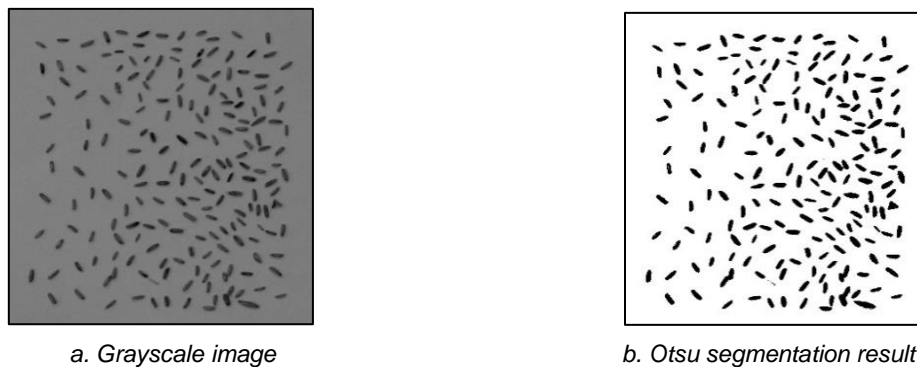


Fig. 2 - Flow chart for oat image processing

Image preprocessing

Oat images are affected by noise such as light source distribution, background selection, image acquisition speed, and external features of the target during acquisition, transmission and storage. In order to remove image noise, this study used median filtering, mean filtering, Wiener filtering, sharpening filtering and other methods to reduce noise for the images. Fig.3a is a grayscale diagram of the oat kernel after the adaptive median filtering process using the median filtering of 7×7 pixel. The experimental results showed that the median filtering could effectively preserve the edge information of the images on basis of eliminating image noise, and the preprocessing effect was good. Therefore, this study used the median filtering results for subsequent experiments.

In order to achieve statistical calculation of oat kernels, the image needed to be segmented to separate the oat kernels from the complex background. The Otsu method was an adaptive threshold segmentation method commonly used in threshold segmentation, and its criterion for determining the optimal threshold is to minimize the variance within each pixel class and maximize the variance between classes. In this study, Otsu method was used to segment the oat kernels, with the segmentation results as shown in Fig.3b. The segmentation threshold was 94 pixels and the separability measure was 0.7673.



a. Grayscale image

b. Otsu segmentation result

Fig.3 - Image segmentation results for oat containing no impurities

As could be seen from Fig.3b, the method was able to separate oat kernels from complex backgrounds.

$$\sigma_b^2(k) = \frac{[m_G p_1(k) - m(k)]^2}{p_1(k)[1 - p_1(k)]} \tag{1}$$

where: m_G represents the average grayscale of the image;
 $m(k)$ represents the cumulative mean;
 $p_1(k)$ represents the probability of occurrence of set C1;

$\sigma_B^2(k)$ represents the variance between classes.

$$\eta_k = \frac{\sigma_B^2(k)}{\sigma_G^2} \quad (2)$$

where: η_k represents a normalized dimensionless matrix

$\sigma_B^2(k)$ represents the variance between classes;

$\sigma_G^2(k)$ represents the total gray scale variance of the image.

Hole filling

In the image processing process, due to the complexity of the environment or the defects of the detection algorithm, the binary images obtained by “binarization” usually had holes, which not only affected the detection effect of the oat kernel, but also destroyed the integrity of the oat kernel profile, which had an adverse effect on the statistical analysis of oat kernels. However, the morphological method could convolve the original image through structural elements to achieve the purpose of removing image noise and eliminating object boundary points (Zhang *et al.*, 2015). Therefore, in this study, the morphological method was used to fill the images with holes, i.e., the images were first etched and then expanded to remove fine debris and separate the slightly adhered oats. After many experiments, it was found that when the structural element of the corrosion operation was set to 5×5 pixel, the short stems could not be eliminated due to the presence of fine impurities in the original image, and the short stems could be removed when the structural element was set to 9×9 pixel. Thus, the effect of fine impurities on the processing results could be eliminated. Fig.4 shows the results of the etching operation and the expansion operation. The filling of the holes was implemented by Equation 3.

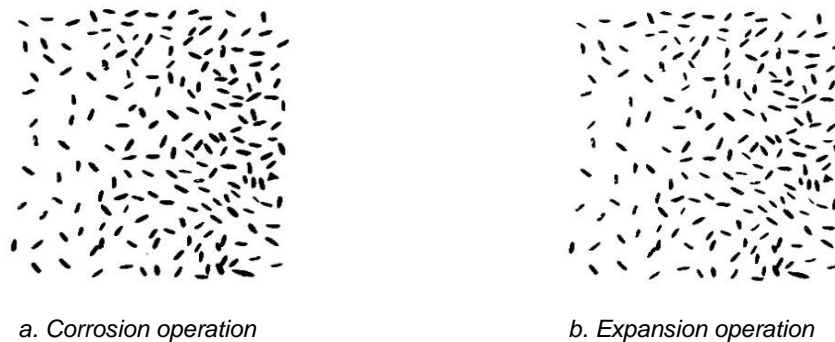


Fig. 4 - Morphological processing results

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad (3)$$

where: X_k represents the k th iteration boundary of set A ;

X_{k-1} represents the $(k-1)$ th iteration boundary of set A ;

B represents the structural element;

A represents collection, with the element of an 8-connected boundary;

A^c represents a complement of A .

STATISTICAL CALCULATION OF THE NUMBER OF OATS BASED ON WATERSHED SEGMENTATION

Segmentation of adhered oats

In grain kernel image processing, due to object adhesion, overlap, etc., often it is impossible to achieve complete and effective segmentation. Through the images of the oat kernel after the threshold segmentation, some of the kernel contacts were dense, there was adhesion and the accuracy of the kernel count was affected. In view of the above situation, the watershed segmentation algorithm introduced the concepts of topography and hydrology into region-based image segmentation, which was suitable for segmentation of adhesion regions. The method regarded the image as a topographic map, the pixel value represented the altitude, the topographical watershed was introduced into the image, and the watershed was used as the image dividing line (Shen *et al.*, 2015). When the rain fell to the side of the watershed, it would flow down the hillside and gather at the lowest point of the valley. Therefore, the valley line was also called the catchment line, and the lowest point at the exit of the valley was called the valley mouth. Since the oat kernel was elliptical, the contour of the cohesive grain had a convex boundary, the binary image of the cohesive grain would form a distinct valley line after the Euclidean distance transformation, and the continuous valley line in

the binary image could be used as the adhesion image segmentation line. In general, there was a problem of over-segmentation by using the watershed segmentation algorithm (Cai *et al.*, 2017). In this study, the marker-based watershed segmentation algorithm was used to segment the oat images, and internal mark and external mark were used to mark them. The internal mark represented the oat kernel and the external mark represented the background. In the segmentation the internal mark was used as low-lying, and the watershed of segmentation result was used as the external mark. The specific implementation process of the algorithm was divided into the following steps.

- (1) Performing Euclidean distance transformation on the binary image to obtain a distance map;
- (2) Calculating the position of a large number of "local minimum regions" in the image, wherein the "local minimum region" had the same brightness value and the brightness values around it were larger than it;
- (3) Calculating the "extended minimum transform" of the image, i.e., expanding the "local minimum region" by a certain threshold, and merging more pixels with similar minimum regions into the minimum region, and taking the "extended local minimum region" as internal tag set;
- (4) Segmenting the image by the general watershed segmentation algorithm, and the watershed of the segmentation result was used as an external mark to obtain the marking result.

Fig.5 shows the segmentation results based on the mark-based watershed segmentation algorithm, where in Fig. 5a was a binary image, and Fig. 5b showed the watershed segmentation result. As can be seen from Fig. 5b, the method could effectively separate the adhered oat kernels.

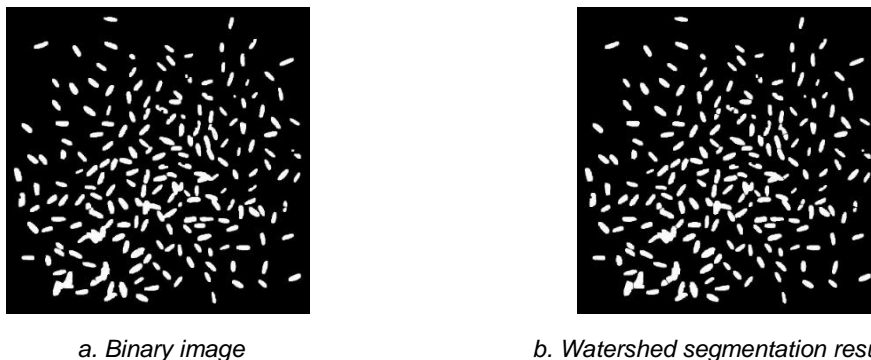
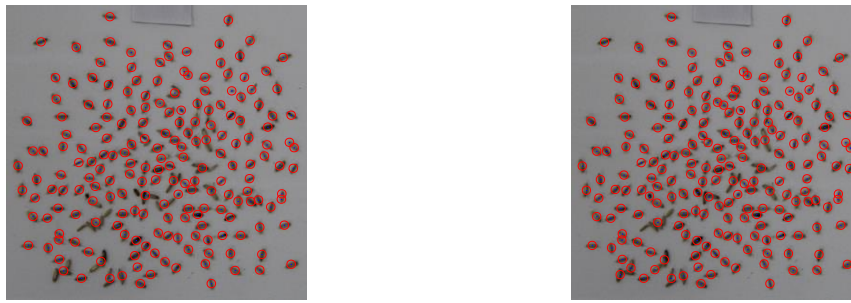


Fig. 5 - Segmentation results based on mark-based watershed segmentation algorithm

Oat kernel count

The image obtained by the watershed segmentation algorithm was a binary one. In the counting statistics of oat kernel, for the image pixels of the oat kernel were continuous, the connectivity was expressed in the binary image. Therefore, this study achieved oat kernel count by finding the number of connected regions in the image. The connected region referred to the region whose value was 0 or 1 in the binary image. After detection, the regions of corresponding pixel shall be determined, and the points with the same pixel value were integrated to form a region, that is, a set of adjacent pixels in the image was detected. After dividing the area in the image, the area, shape and other parameters of the region could be calculated to provide a basis for analyzing the classification and measurement of oat kernels.

The connected area of the image was obtained by scanning the image. According to different neighbouring relationship of the pixels, the connected regions were generally divided into 4-neighbour regions and 8-neighbour regions. The 4- neighbour regions were scanned in four directions of up, down, left, and right of the detection points, and the 8-neighbour areas were scanned in eight directions of up, down, left, and right as well as four diagonal directions of the detection points. Generally, the number of neighbour regions was no less than that of 8-neighbor regions. In this study, the 8-neighbor region analysis was used to mark oat kernels and calculate the area and number of regions in each region. Fig.6 was a statistical result of the image connectivity area of 5g oat kernel image, in which the actual number of oat kernels was 207. In Fig.6a is showed the statistical result of Otsu segmentation, in which the number of oats was 195, while in Fig.6b is showed the statistical result of the division of the adhered oat kernels by watershed, and the number of oats was 204. As can be seen from the figure, the method was able to mark the oat kernels in the image well, and the oat kernels after the watershed segmentation could also be well marked, indicating that the number of oat kernels could be obtained by the image analysis method. Through statistical analysis, the average area of oat kernels was 295 pixels.



a) Otsu segmentation statistics

b) Watershed segmentation statistics

Fig. 6 – Results of analysis of connected regions

Separation of impurities by area method

According to the above statistical results, the area parameters of the oat kernels were obtained. As the actual oats were obtained, they contained impurities such as stalks, twigs, hulls, leaves, etc., and the colour of the impurities was close to the colour of the oats, they were different mainly in shape and size, therefore, the area method was adopted to identify and separate the impurities, and the area threshold T was selected. When $T > T_{\text{mean}}$, the pixel value of the point was marked as the background, that is, the impurities were removed. In this study, after several experiments, the area threshold was selected to be 300 pixels and 600 pixels to remove impurities. The selection of the threshold affects the recognition rate of impurities. Fig.7 showed the result of removal of impurities by the area method. When T was 600 pixels, the experimental result showed that impurities was better removed and while the basic information of oats was retained. When T was 300 pixels, as shown in Fig.7c, while removing impurities, some oats were also removed, and there was over-segmentation. Therefore, in this paper, it was finally determined the threshold value would be 600 pixels, and the oat image after removing impurities was obtained.



a. Original image with impurities contained

b. Area threshold 600

c. Area threshold 300

Fig. 7 – Results of impurity removal by area method

RESULTS AND DISCUSSION

Analysis of loss rate of cleaning of oats without impurities

The loss rate of oat cleaning was expressed by the ratio of mass loss after the sample going through the cleaning machine and the mass before operation, as stated in Formula 4.

$$L = \frac{\Delta w}{w_1} \times 100\% \quad (4)$$

where: w_1 represents the sieve mixture, [g];
 Δw represents change in weight, [g];
 L represents the loss rate, [%].

w_1 was obtained by using an electronic balance, while w_2 was calculated by referring to the thousand-grain weight of oats. The number of oat kernels shall be obtained by the algorithm, and the calculation of w_2 could be based on the following Formula 5.

$$W_2 = \frac{ax}{1000} \quad (5)$$

where: w_2 represents the grain under the sieve, [g];
 a represents thousand-grain weight of oats, [g];
 x represents the number of oat kernels obtained by the algorithm, [grain].

In this study, the reference value of sample a was 25.5g, and the test data obtained in group a were substituted into Formula 4 and Formula 5, respectively, and the results were shown in Table 1.

Algorithm statistical number value and loss rate of cleaning

Table 1

| W_1 [g] | Average number [grain] | Algorithm result [grain] | Recognition rate [%] | W_2 [g] | Δw [g] | L [%] |
|-----------|------------------------|--------------------------|----------------------|-----------|----------------|-------|
| 5 | 194 | 192 | 98.97 | 4.90 | 0.10 | 2.08 |
| 8 | 308 | 231 | 75.00 | 5.89 | 2.11 | 26.37 |
| 10 | 382 | 239 | 62.57 | 6.09 | 3.91 | 39.06 |
| 12 | 459 | 264 | 79.30 | 6.73 | 5.27 | 43.90 |
| 15 | 582 | 261 | 44.84 | 6.66 | 8.34 | 55.63 |
| 18 | 708 | 225 | 32.34 | 5.74 | 12.26 | 68.13 |
| 20 | 775 | 313 | 40.38 | 7.98 | 12.02 | 60.09 |

With the data in Table 1, it could be seen that the 5 g-sample recognition rate of the algorithm reached 98.97%, and the calculated loss rate of cleaning was 2.08%, which was in line with national standards, while the recognition rate of other groups was generally lower. The 18 g-sample was with the lowest recognition rate of only 32.34%, with a corresponding cleaning loss rate of 68.13%. Analysis of the image processing process showed that the oat image obtained from the 5 g-sample had less adhesion of oat kernels, and no overlap or cover was there, also the algorithm statistics were accurate, and the statistical results of the algorithm were accurate. The recognition rate of samples from 8g to 20g was generally low, with the recognition rate basically negatively correlated with the weight of oats, and calculation result of loss rate far exceeded the national standard, for which the reason was that as the sample quality increased, more large-area adhesion and overlap occurred, but the recognition ability of the algorithm was limited in face of these conditions, so the statistically calculated loss rate of cleaning could not reflect the actual loss.

Analysis of loss rate of oats containing impurities

It can be seen in Table 1 that the identification rate of other oat samples other than the 5 g-sample was generally low, and the calculation of the loss rate could not reflect the actual cleaning loss. Therefore, in the experiment with impurities, other groups were discarded, leaving only the 5g group, where the adhesion region well handled by the watershed segmentation algorithm. Short stalks and some fine impurities were separated by area method. After many experiments, it was determined that when the threshold value was 600 pixels, the short stalks and fine impurities could be well removed, while the basic information of oats could be well preserved. In the 5g group test, the w_2 was calculated to be 4.87g, and the L was 2.60%, which was slightly higher than the no-hybrid statistics and basically met the standard. The results were shown in Table 2.

Table 2

Statistical calculation of the number of oats containing impurities and the loss rate of cleaning

| W_1 [g] | Average number [grain] | Algorithm result [grain] | Recognition rate [%] | W_2 [g] | Δw [g] | L [%] |
|-----------|------------------------|--------------------------|----------------------|-----------|----------------|-------|
| 5 | 194 | 191 | 98.97 | 4.87 | 0.13 | 2.60 |

CONCLUSIONS

(1) Based on the content of this experiment, the watershed segmentation algorithm was effective in calculating the loss rate of oat in its cleaning process, also the loss rate could be calculated by directly acquiring images in the cleaning process. Currently this method was limited to small masses.

(2) To obtain the impurity-free image, the Otsu algorithm shall be adopted to pre-segment the grain and the background image, to get a binary image in which "holes" might appear. Then the morphological opening algorithm could be used to fill the image with holes, so as to remove fine debris and separate the slightly adhered oats.

(3) In this study, the condition of the oat kernels containing impurities was relatively simple, and the impurities were stalks having a large difference in shape from oat. In actual situations, a mixture of various impurities might be there. For example, the impurities might also contained hulls or dust residue similar to oats in shape. If the fixed threshold of this paper could not achieve the expected effect, in the future, it is desirable to find a self-adapting threshold to remove different impurities.

(4) The calculation of the cleaning loss rate was closely related to the quality of image obtained. Designing a suitable buffering facility at the receiving position of the cleaning machine could minimize the adhesion and overlap regions to the maximum, so that the calculation of the loss rate could be more accurate.

(5) In the case of increased sample quality, large adhesion-and-overlapping regions might occur. For these regions, the number count and the estimation of loss rate of cleaning might be distorted, and the seriousness of distortion is negatively correlated with quality. It is the content of following research of this research group as to how to solve the problem of identifying the number of grains in large adhesion-and-overlapping regions and estimating their weight.

ACKNOWLEDGEMENTS

This research, titled 'Estimation of Loss Rate of Oats Cleaning Based on Watershed Segmentation', was funded by the National Key Research and Development Plan of China (2016YFD0701801). The authors are grateful and honoured to have obtained support from the Key Laboratory of Biomechanics.

REFERENCES

- [1] Cai Q., Liu Y.Q., Cao J., Li H.S., Du J.P., (2017), A Watershed Image Segmentation Algorithm Based on Self-adaptive Marking and Interregional Affinity Propagation Clustering. *Acta Electronica Sinica*, Vol.45, Issue 8, pp.1911-1918, Beijing/P.R.C.;
- [2] Diego I.P., Rafael R., (2018), Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, Vol.153, pp.69-81, Oxford/England;
- [3] Eric I.W., (2018), *Grain loss sensor array for crop harvesting machine*. Patent No. 20180053067A1, USA;
- [4] Gao J.M., Zhang G., Yu L., Li Y.B., (2011), Chaos detection of grain impact at combine cleaning loss sensor. *Transactions of the Chinese Society of Agricultural Engineering*, Vol.27, Issue 9, pp. 22-27, Beijing/P.R.C.;
- [5] Liang Z.W., Li Y.M., Xu L.Z., Tang Z., (2015), Monitoring Mathematical Model of Grain Cleaning Losses on Longitudinal-axial Flow Combine Harvester. *Transactions of the Chinese Society for Agricultural Machinery*, Vol.46, Issue 1, pp.106-111, Beijing/P.R.C.
- [6] Liang Z.W., Li Y.M., Xu L.Z., Zhao Z., Tang Z., (2017), Optimum design of an array structure for the grain loss sensor to upgrade its resolution for harvesting rice in a combine harvester. *Biosystems Engineering*, Vol.157, pp. 24-34, San Diego/USA;
- [7] Mao H.P., Liu W., Han L.H., Zhang X.D., (2012), Design of intelligent grain cleaning losses monitor based on symmetry sensors. *Transactions of the Chinese Society of Agricultural Engineering*, Vol.28, Issue 7, pp. 34–39, Beijing/P.R.C.;
- [8] Shen X.J., Wu X.Y., Han D.J., (2015), Survey of Research on Watershed Segmentation Algorithms. *Computer Engineering*, Vol.41, Issue 10, pp.26-30, Shanghai/P.R.C.;
- [9] Tang Q., Wu C.Y., Wang S.Z., Zhang C.W., Sun Y.F., (2013), Research Advances and Prospects in Cleaning Device of Grain. *Journal of Agricultural Mechanization Research*, Vol.12, pp.225-228, Harbin/P.R.C.;
- [10] Uryi I.M., Aleksei V. P., Aleksandr Y. S., Ivan A. K., Dmitry N. P., Rinat N. S., (2018), Computer vision system: a tool for evaluating the quality of wheat in a grain tank, *Proceeding of the 10th International Conference on Machine Vision*, Vol.10696, pp.451-457, Vienna/ Austria;
- [11] Vithu P., Moses J.A., (2016), Machine vision system for food grain quality evaluation: A review. *Trends in Food Science & Technology*, Vol.56, pp.13-20, London/England;
- [12] Wei C., Li M., Yu L.J., Zhu L., (2016), Research on Loss Monitoring of Grain Cleaning in Combine Harvester Based on Kalman Filter. *Journal of Agricultural Mechanization Research*, Vol.3, pp.23-27, Harbin/P.R.C.;
- [13] Zhang G., Gao J.M., Yu L., (2012), Chaos Detect System of Cleaning Loss Based on ARM9. *Journal of Agricultural Mechanization Research*, Vol.3, pp.159-161+216, Harbin/P.R.C.;
- [14] Zhang Y.H., Li Y., Wang F.Q., Zhang Z.X., (2015), A Watershed Segmentation Algorithm Based On Morphological Reconstruction and Maxima Mark. *Computing Technology and Automation*, Vol.2, pp.78-81, Changsha/P.R.C.;
- [15] Zhou X.L., Zhu R.X., Zhou X.J., Tang Y., (2010), The Monitoring System for Cleaning Loss of the Grain Combine Harvester Based on Sensor Technology. *Journal of Agricultural Mechanization Research*, Vol.12, pp.85-87, Harbin/P.R.C.