Detection of Maize Kernels Breakage Rate Based on K-means Clustering

Liang Yang1, 2, 3, a), Zhuo Wang1, 2, 3, Lei Gao1, 2 and Xiaoping Bai1, 2

1Key Laboratory of Networked Control System, Chinese Academy of Sciences, Shenyang 110016, China.
2Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China.
3University of Chinese Academy of Sciences, Beijing 100049, China.

a)yangliangsia@sina.cn

Abstract. In order to optimize the recognition accuracy of maize kernels breakage detection and improve the detection efficiency of maize kernels breakage, this paper using computer vision technology and detecting of the maize kernels breakage based on K-means clustering algorithm. First, the collected RGB images are converted into Lab images, then the original images clarity evaluation are evaluated by the energy function of Sobel 8 gradient. Finally, the detection of maize kernels breakage using different pixel acquisition equipments and different shooting angles. In this paper, the broken maize kernels are identified by the color difference between integrity kernels and broken kernels. The original images clarity evaluation and different shooting angles are taken to verify that the clarity and shooting angles of the images have a direct influence on the feature extraction. The results show that K-means clustering algorithm can distinguish the broken maize kernels effectively.

Key words: Computer vision; K-means clustering; detection of the maize kernels breakage; energy function of gradient; image clarity evaluation.

INTRODUCTION

The broken rate of maize kernels is an important index to measure the performance of maize harvester. Therefore, it is of great research significance to improve the detection accuracy of maize kernels breakage, which is an important way to improve harvest performance.

In China, Yan Zhao [1] based on computer vision technology, can effectively identify the moldy maize seeds to ensure the storage quality of maize seeds. Hong Zhou [2] based on computer vision technology and image processing technology to extract the maize contour, using the fuzzy pattern recognition to replace the traditional appraisal method of maize seeds. Xiaomei Yan [3] put forward a new method of maize seeds purity identification. The method was proposed based on color extracted from the images of both the maize crown and the maize side for improving the purity of maize seed.

In other countries, some researchers focused on studying integrity of maize kernels. Zayas I [4] selected 7 morphological parameters with better contributions to be identified from 12 morphological parameters and analyzed the 7 morphological parameters by establishing Mahalanobis distance discriminant function. The purpose of the method is that it distinguished between the integrity maize kernels and the broken maize kernels. Ding K [5] used continuous symmetric exponential, cervical curvature exponential and the exponential of the radius to distinguish the contour integrity between the maize kernels and broken maize kernels. This method can effectively identify the broken maize kernels from the overall. Liao K [6] used plurality parameters to describe the image feature of the maize kernels and combined computer vision technology and neural network algorithm to distinguish the integrity maize kernels and the broken maize kernels.

Based on the above document analysis, this paper presents a clustering algorithm of K-means for the detection on maize kernels breakage.
K-MEANS CLUSTERING ALGORITHM AND IMAGE CLARITY EVALUATION

K-means Clustering Algorithm

K-means clustering algorithm is simple and rapid. The algorithm basic idea is given the spacing of \( n \) samples that have known their pixel sizes, and picking \( k \) cluster centers, randomly. Then evaluation the degree of similarity between the samples based on Euclidean distance. Every element is allocated to the cluster which provides the smallest Euclidean distance to the centroid of each cluster. And the average value of the samples categorized into each class is calculated and updates the center of each classes until the square error function is stable at the minimum value.

The Euclidean distance is the sum of square of the distances from each data point to the center of the sample, is also called the objective function. Supposing the objects collection is \([7-8]\):

\[
N = \{x_1, x_2, x_3, ..., x_n\}, \quad x_i = (x_{i1}, x_{i2}, x_{i3}, ..., x_{in})
\]

The Euclidean distance formulate between the sample \( x_i \) and the sample \( x_j \) is:

\[
d(x_i, x_j) = [(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2 + ... + (x_{in} - x_{jn})^2]
\]  

(1)

The square error function is:

\[
I_c = \sum_{i=1}^{K} \sum_{j=1}^{t_i} \| x_j - n_i \|^2
\]

(2)

Where \( K \) is the number of clusters to be clustered. \( t_i \) is the number of samples for the \( i \) class. \( n_i \) is the mean of the samples in the \( i \) class.

Kernel Recognition Based on Lab Color Model

Before using K-means clustering algorithm, the image from RGB color model to Lab color model, the method is as follows [9]:

\[
N_L = 0.2126 \times N_R + 0.7152 \times N_G + 0.0722 \times N_B \\
N_a = 1.479 \times (0.2213 \times N_R - 0.3390 \times N_G + 0.1177 \times N_B) + 128 \\
N_b = 0.6245 \times (0.1949 \times N_R - 0.6057 \times N_G - 0.8006 \times N_B) + 128
\]

(3)

Where \( N_L \) is the \( L \) value of a pixel in a Lab image. \( N_a \) is the \( a \) value of a pixel in a Lab image. \( N_b \) is the \( b \) value of a pixel in a Lab image. \( N_R \) is the \( R \) value of a pixel in a RGB image. \( N_G \) is the \( G \) value of a pixel in a RGB image. \( N_B \) is the \( B \) value of a pixel in a RGB image.

Image Clarity Evaluation

Image clarity evaluation reflects the property of the image clarity and clarity is an important property of an image. Reasonable clarity can provide the benefit of subsequent processing of the image and identification of the image characteristics. At the same time, the image clarity evaluation function is also the key to realize automatic focusing in digital image acquisition system. The focusing performance depends on the accuracy and the real-time of the image clarity evaluation function [10-11]. In the energy function of gradient, the larger of the energy values of the edge gradient, the clearer the image and the greater difference the value of the evaluation function between the target region and the background region. The formula of energy function of gradient is:
\[ E(i, j) = |I(i+1, j) - I(i, j)|^2 + |I(i, j+1) - I(i, j)|^2 \]  \hspace{1cm} (4)

Where \( E(i, j) \) is the evaluation function of point \((i, j)\), \( I(i, j) \) is the gray value of point \((i, j)\).

In this paper used the energy function of Sobel 8 gradient to evaluate the clarity of the acquired images, and its clarity evaluation value is the sum of the square of the edge pixel gradient.

**TEST DESIGN**

The integrity maize kernels were mixed with the broken maize kernels and the mixed maize kernels and they were used as the experimental sample. In the experiment using two devices with different pixels. The valid pixel of device 1 is 12.2 million pixels and the valid pixel of device 2 is 800 million pixels. Colleting the experimental from different shooting angles sample by respectively using the two devices, their clarity evaluation value are different. After sampling, using Matlab2014a evaluated the pictures clarity and processed the pictures by using K-means clustering algorithm.

The original images were collected as shown in the figure:

![Experimental sample](image1)

**FIGURE 1.** Experimental sample.

![Image information acquired by the device 1](image2)

**FIGURE 2.** Image information acquired by the device 1.

![Image information acquired by the device 2](image3)

**FIGURE 3.** Image information acquired by the device 2.
Using the energy function of Sobel 8 gradient evaluated the clarity of those images, and L shows the values. The larger value of the L indicates the higher the image clarity. Results:

![Images of devices](image1.png)

(a) 4.1656e+10  (b) 4.8531e+10  (c) 3.9345e+10

**FIGURE 4.** Show the image clarity evaluation values of the device 1.

![Images of devices](image2.png)

(a) 1.9330e+10  (b) 1.5057e+10  (c) 2.1002e+10

**FIGURE 5.** Show the image clarity evaluation values of the device 2.

To identify the broken maize kernels in the image by using K-means clustering algorithm. The recognition rate is the ratio of the broken maize kernels to the total maize kernels. Results of K-means clustering:

![Images of devices](image3.png)

**FIGURE 6.** Device 1 K-means clustering result.

![Images of devices](image4.png)

**FIGURE 7.** Device 2 K-means clustering result.
TABLE 1. Experimental results.

<table>
<thead>
<tr>
<th>Filming angle</th>
<th>Device 1 recognition rate</th>
<th>Device 2 recognition rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>-45°</td>
<td>78%</td>
<td>72%</td>
</tr>
<tr>
<td>0°</td>
<td>90%</td>
<td>83%</td>
</tr>
<tr>
<td>45°</td>
<td>83%</td>
<td>70%</td>
</tr>
</tbody>
</table>

CONCLUSION

In this experiment, which is based on K-means clustering algorithm, illustrated the detection of maize kernels breakage is related to the image clarity and shooting angles. At the same time, it was verified that the K-means clustering algorithm was feasible in the detection of maize kernels breakage. But K-means clustering algorithm has a lower recognition rate at different clarity and different angles. Therefore the algorithm need to be further optimized and the maize kernels size, plumpness need to be introduced to this experiment. Meanwhile, the optimization of K-means clustering algorithm can effectively identify the broken maize kernels under the condition of low image clarity, more noise and less effective area of broken characteristics, which will be a further step in studying the detection of maize kernels breakage.

ACKNOWLEDGMENTS

This work was financially supported by Science and Technology Planning Project of Liaoning Province (2015104005), China fund.

REFERENCES