ICIC Express Letters
Volume **10**, Number **6**, June **2016**

ICIC International ©2016 ISSN 1881-803X
pp. 1273–1278

A SIMULATION OF NATURAL WOLFBERRY RECOGNITION BASED ON COLOR CLASSIFICATION

Cheng-Hung Chuang, Jui-Chi Chen and Yung-Chen Chou

Department of Computer Science and Information Engineering
Asia University
No. 500, Liufeng Rd., Wufeng, Taichung 41354, Taiwan
{ chchuang; rikki; yungchen }@asia.edu.tw

Received December 2015; accepted March 2016

Abstract. *This paper presents a computer simulation study of natural wolfberry recognition based on color feature. We performed a computer simulation of the recognition between natural and sulfur-smoked wolfberries by using color classification for food safety. First, two types of images of the natural and sulfur-smoked wolfberries were collected in the Internet. These images were converted from RGB to CIELab color space and the color feature vectors were obtained by the combination of chromatic components. Then the color vectors of these two types of images were classified by the k-means clustering algorithm in the training process. By means of the color classification, the recognition between natural and sulfur-smoked wolfberries is performed. Simulation results showed that the proposed method provides a precision rate of about 92% in the collected image dataset.*

Keywords: Chinese wolfberry, Goji berry, Image processing, Pattern recognition, Food safety

1. **Introduction.** Wolfberry, also known as Chinese wolfberry, goji berry, or simply called goji, is the fruit of lycium barbarum. It is a kind of drug that is commonly used in traditional Chinese medicine [1] and is also one of the popular functional foods that is usually used to make tea or eat with food. In order to have better effect, many people like to choose bright red wolfberries at the time of purchasing. However, bright red wolfberries are suspected to be caused by sulfur fumigation.

The dried wolfberries smoked by sulfur can not only prevent pest and extend the shelf life, but also present the bright red color that makes them sell better. Through the sulfur fumigation in food, the sulfur dioxide is produced. It can have an effect of bleaching food and providing bright color on food. If one has eaten the sulfur-smoked wolfberries for a long time, his health will be greatly harmed.

To simply solve the problem of food safety of wolfberries, we can select dark red wolfberries instead of bright red wolfberries. However, it is a challenge to classify wolfberries by human eyes. Therefore, we want to develop an automatic recognition system of natural wolfberries based on color feature. We can capture images of wolfberries and match the color feature through the recognition system to identify natural or sulfur-smoked wolfberries.

In [2], an optimal probabilistic neural network based on a two-step numerical procedure is proposed to classify food color. The proposed method is applied to classifying French fries into three different color classes, i.e., light, normal, and dark. In [3], the image analysis using the intensity of darkening during frying and artificial neural network is used to classify the color of fried potato chips. The proposed model is used to estimate the quality of potatoes in chips industry. In [4], a color machine vision algorithm that consists of two components is proposed to solve the problem of color variability. The first component creates an artificial color contrast as a pre-filter that aims at highlighting the target while suppressing its surroundings. The second component utilizes the principal component analysis technique to characterize target features in color space from a set of training data so that the color classification can be performed accurately and efficiently. These researches indicated that the color information is useful in vision-based feature detection, particularly for food processing applications.

The color space plays an important role in the recognition system. CIELab is a color space specified by the International Commission on Illumination (CIE). It is the most exact means of representing color and is created to serve as a device-independent model to be used as a reference [5]. In literature, there is a lot of research work using the CIELab color space. Their experimental results show that CIELab is superior to RGB color model for the representation of color feature [6-8].

In [9], the authors proposed a lip detection method using CIELab color space. Then a fuzzy C-means clustering classification method is applied to classifying image pixels and finally detecting red lips using the elliptical shape and the Gaussian mask. In [10], the authors proposed a lesion detection method in digital color fundus photographs. They used the region growing algorithm combined with k-nearest neighbor classifier to successfully identify the red lesions. In [11], the authors used a color compensation chart to be a reference standard for image color correction. In [12], the authors proposed a 3-dimensional thin-plate spline interpolation method for color correction.

In our proposed simulation, we collected two types of wolfberry images. The different color features of the chromatic components in the two types of images are extracted. Then the recognition between natural and sulfur-smoked wolfberries is performed by the color classification. In Section 2, we described the proposed simulation method. Section 3 covers some simulation results to demonstrate the performance of the proposed method. Finally, Section 4 concludes this study.

2. The Proposed Method.

In this paper, the image processing and pattern recognition technology are used to develop a natural wolfberry recognition system to achieve the goal of food safety. In the proposed method, the main steps are image segmentation, color space conversion, color classification, and color feature matching. Figure 1 shows the block diagram of the proposed method. The sample images are first segmented into regions of interest (ROIs), i.e., regions of wolfberries, and converted into CIELab color space. The chromatic components of ROIs are taken as color feature vectors. In the training process, the color feature vectors of all pixels of ROIs are classified into two groups according to their different types of natural or sulfur-smoked wolfberries. In the test process, test images are also segmented and converted into CIELab color space. Then based on the color classification, chromatic components of ROIs of test images are matched to two different groups. The percentages of chromatic components of wolfberries that belong to two kinds of groups are computed respectively.

2.1. Image segmentation.

Image segmentation is an important step in the recognition method to partition wolfberries in images. In this step, we assumed that the red wolfberries are put on a white paper, i.e., the target is red and the background is white. We split

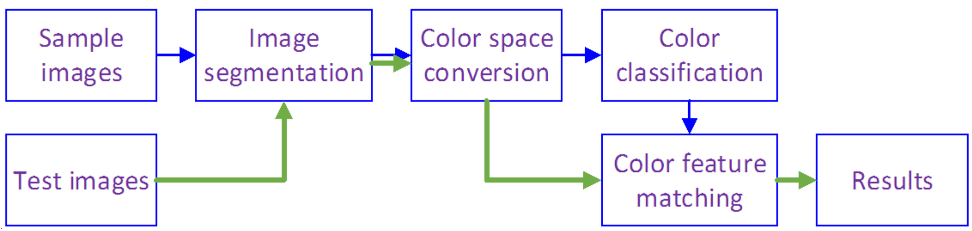

Figure 1. Block diagram of the proposed method

the RGB color image into three channel images, i.e., red, green, and blue. Because the blue channel of color images gives the highest contrast between target and background, this channel is used for the extraction of the wolfberry regions. Then the segmentation of the blue channel image is achieved by the Otsu thresholding method [13]. We then obtain a binary image that red wolfberries are segmented but with some small holes. To solve this problem, the algorithm based on morphological reconstruction [14] is used to fill in the holes to obtain proper wolfberries. Figure 2 shows a sample result of the segmentation method.

Figure 2. Sample result of the segmentation method

2.2. **Color space conversion.** The CIELab model is a color space with an L component for lightness dimension, and a and b components for the color dimensions. The a component indicates the color falling along the red-green axis, and the b component specifies the color falling along the blue-yellow axis. Because the red-green and blue-yellow opponent channels are computed as differences of lightness transformations of cone responses, CIELab is a chromatic value color space. The color feature vector of each image pixel is obtained by the combination of chromatic components.

2.3. **Color classification.** After the image segmentation and CIELab color space conversion, we can obtain the color feature vector of every pixel on the wolfberry regions. Then we used the k-means clustering algorithm to classify color feature vectors into two groups of natural and sulfur-smoked wolfberries in the training process. We assumed that the color feature vectors of all pixels on the ROIs of sample images are a set of observations (x_1, x_2, $\ldots$, x_n), where each observation is a two dimensional color vector, and the k-means clustering aims to partition the n observations into two sets $S = \{S_1, S_2\}$ so as to minimize the within-cluster sum of squares. In other words, its objective is to find the argument of the minimum as follows.

$$\arg\min_{S} \sum_{i=1}^{2} \sum_{x \in S_i} \|x - \mu_i\|^2 \tag{1}$$

where μ_i indicates the mean of the set S_i, $i = 1, 2$.

2.4. **Color feature matching.** In the test process, a test image is segmented into ROIs, where ROIs correspond to regions of wolfberries, and converted into CIELab color space. The color feature vectors of all pixels on the ROIs are obtained. We assumed that the color feature vectors of all pixels on the ROIs of a test image are a set of observations $X = (x_1, x_2, \ldots, x_m)$, the color feature matching method matches the m observations into two sets $S = \{S_1, S_2\}$ so as to compute the percentage of observations belonging to each set of S.

3. **Simulation Results.** In this section, we present the computer simulation results of the proposed natural wolfberry recognition system.

3.1. **Collection of wolfberry images.** We first collect wolfberry images in the Internet. According to the statements on the websites, the wolfberry images are categorized into two kinds of types, i.e., natural and sulfur-smoked wolfberry images. There are 24 wolfberry images in which 12 images are natural and 12 images are sulfur-smoked. These images are segmented to obtain the regions of wolfberries. Figure 3 shows some sample wolfberry images in which the natural wolfberry images are in left column and the sulfur-smoked wolfberry images are in right column.

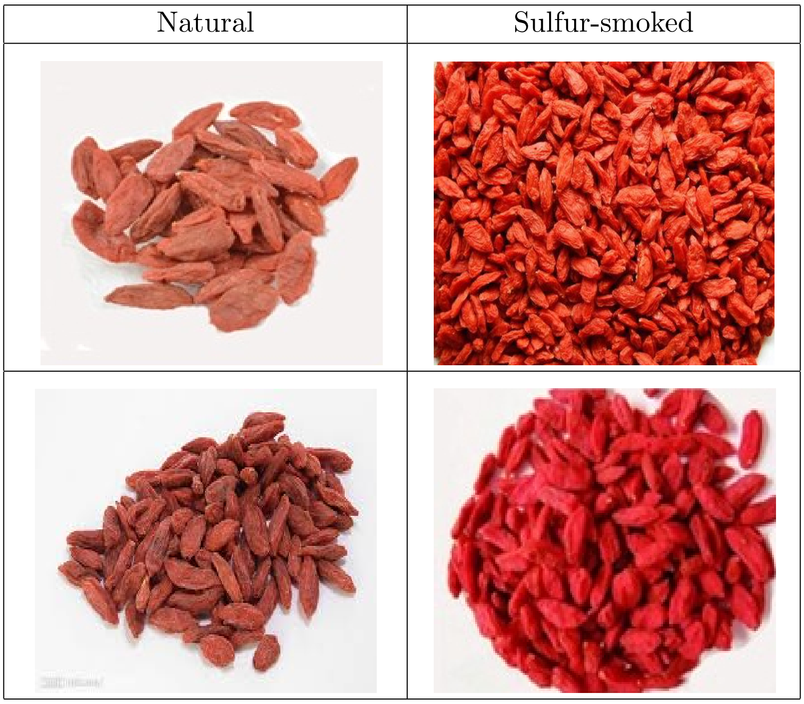

FIGURE 3. Sample wolfberry images

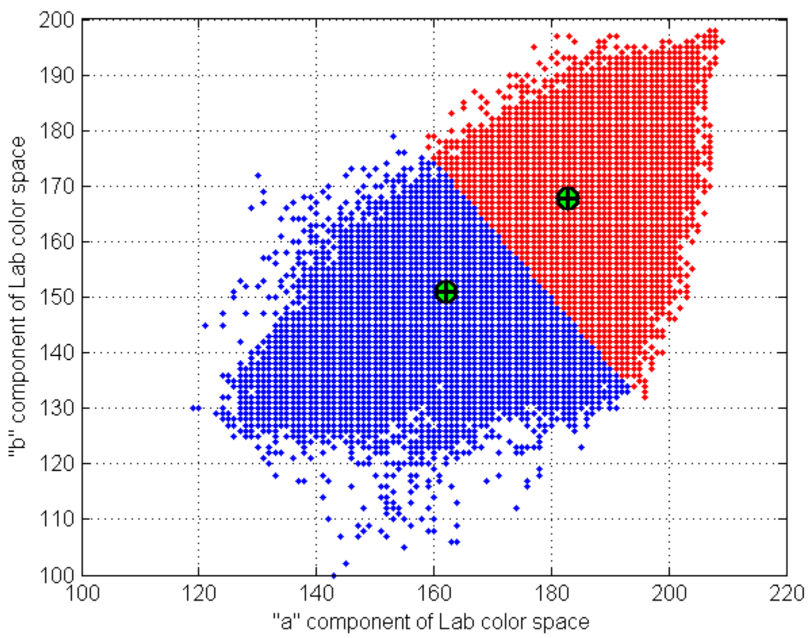

FIGURE 4. Result of the k-means clustering

3.2. **Simulation of color classification.** In the color classification process, we selected a half of the collected wolfberry images in which 6 images are natural and 6 images are sulfur-smoked. The k-means clustering algorithm is used to classify color feature vectors into two groups of natural and sulfur-smoked wolfberries. Figure 4 shows the result of the k-means clustering, where the blue color (the bottom left side) represents the type of natural wolfberries while the red color (the top right side) indicates the type of sulfur-smoked wolfberries. The two green circles are the center points of the two groups, respectively. Then the two center points are used to classify color feature vectors of all pixels on the ROIs of the test image.

3.3. **Simulation of color feature matching.** In the color feature matching process, we examined the remaining collected wolfberry images in which 6 images are natural (No. 1∼6) and 6 images are sulfur-smoked (No. 7∼12). We extracted the color feature vectors of all pixels on the ROIs of a test image. The percentage of color feature vectors

TABLE 1. Percentages of test images belonging to two types of wolfberries

No. of test images	Natural (%)	Sulfur-smoked (%)	Type of matching
No. 1	72.60	27.40	Natural
No. 2	29.87	70.13	Sulfur-smoked
No. 3	71.85	28.15	Natural
No. 4	80.68	19.32	Natural
No. 5	96.59	3.41	Natural
No. 6	58.56	41.44	Natural
No. 7	46.23	53.77	Sulfur-smoked
No. 8	3.47	96.53	Sulfur-smoked
No. 9	26.45	73.55	Sulfur-smoked
No. 10	36.30	63.70	Sulfur-smoked
No. 11	29.40	70.60	Sulfur-smoked
No. 12	16.63	83.37	Sulfur-smoked

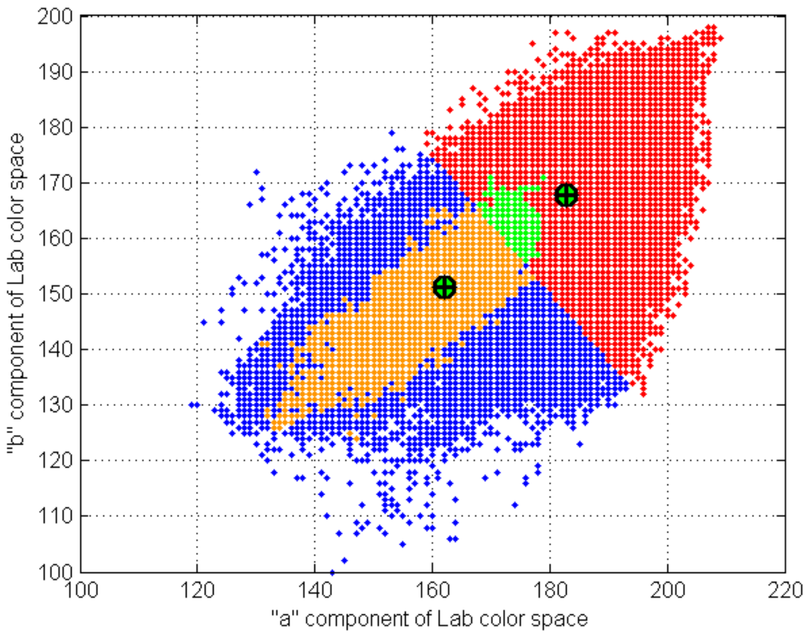

FIGURE 5. Result of color feature matching of image No. 5

belonging to which type of wolfberries is computed. Table 1 shows the percentages belonging to natural or sulfur-smoked. The test image No. 2 is the only image which has the wrong type. The precision is about 92%. Figure 5 shows the result of color feature matching of the test image No. 5, where the percentages belonging to natural and sulfur-smoked type are 96.59% (the central part located in the bottom left side) and 3.41% (the central part located in the top right side), respectively.

4. **Conclusions.** We have applied the color classification method to a natural wolfberry recognition system for food safety. The simulation results show that it is possible to use the system to automatically recognize natural wolfberries. This study is in its infancy. In future work, we will use real wolfberries to validate the feasibility of the method.

Acknowledgment. This work was supported by the Ministry of Science and Technology of Taiwan under Grant No. MOST 103-2221-E-468-008-MY2.

REFERENCES

[1] S. Li, *Compendium of Materia Medica*, Ming Dynasty, AD 1368-1644.
[2] Y. Chtioui, S. Panigrahi and R. Marsh, Conjugate gradient and approximate Newton methods for an optimal probabilistic neural network for food color classification, *Optical Engineering*, vol.37, no.11, pp.3015-3023, 1998.
[3] T. Marique, A. Kharoubi, P. Bauffe and C. Ducattillon, Modeling of fried potato chips color classification using image analysis and artificial neural network, *Journal of Food Science*, vol.68, no.7, pp.2263-2266, 2003.
[4] K. M. Lee, Q. Li and W. Daley, Effects of classification methods on color-based feature detection with food processing applications, *IEEE Trans. Automation Science and Engineering*, vol.4, no.1, pp.40-51, 2007.
[5] Lab Color Space, *Wikipedia*, http://en.wikipedia.org/wiki/Lab_color_space.
[6] M. Corbala'n, M. S. Milla'n and M. J. Yzuel, Color pattern recognition with CIELAB coordinates, *Optical Engineering*, vol.41, no.1, pp.130-138, 2002.
[7] M. S. Alam, S. F. Goh and S. Dacharaju, Three-dimensional color pattern recognition using fringe-adjusted joint transform correlation with CIELab coordinates, *IEEE Trans. Instrumentation and Measurement*, vol.59, no.8, pp.2176-2184, 2010.
[8] S. M. Lajevardi and H. R. Wu, Facial expression recognition in perceptual color space, *IEEE Trans. Image Processing*, vol.21, no.8, pp.3721-3733, 2012.
[9] R. Rohani, S. Alizadeh, F. Sobhanmanesh and R. Boostani, Lip segmentation in color images, *Proc. of International Conference on Innovations in Information Technology*, pp.747-750, 2008.
[10] M. Niemeijer, B. van Ginneken, J. Staal, M. S. A. Suttorp-Schulten and M. D. Abràmoff, Automatic detection of red lesions in digital color fundus photographs, *IEEE Trans. Medical Imaging*, vol.24, no.5, pp.584-592, 2005.
[11] S.-H. Lee, S.-W. Um and J.-S. Choi, Design and implementation of color reproduction system for images captured by digital camera, *Proc. of International Conference on Consumer Electronics*, 2008.
[12] P. Menesatti, C. Angelini, F. Pallottino, F. Antonucci, J. Aguzzi and C. Costa, RGB color calibration for quantitative image analysis: The 3D thin-plate spline warping approach, *Sensors*, vol.12, no.6, pp.7063-7079, 2012.
[13] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Systems, Man, and Cybernetics*, vol.9, no.1, pp.62-66, 1979.
[14] P. Soille, *Morphological Image Analysis: Principles and Applications*, Springer-Verlag, pp.173-174, 1999.