Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Ripe Tomato Extraction For A Harvesting Robotic System

Hongpeng Yin, Yi Chai
College of Automation
Chongqing University
Chongqing City, China
yinhongpeng@gmail.com

Simon X. Yang, Gauri S. Mittal
School of Engineering
University of Guelph
Guelph, Ontario, Canada
syang@uoguelph.ca

***Abstract*—A robotic system for harvesting tomatoes in greenhouses is designed. Effective recognition of ripen tomatoes from complex background is the key technology of the harvesting robotic system. In this work, the color feature of ripen tomatoes is employed. The ripen tomato is segmented by K-means clustering using the L*a*b* color space. To extract a single integrity ripen tomato, mathematical morphology method is used to denoise and handle the situations of tomato overlapping and shelter. Experimental results show the effectiveness of the proposed method.**

***Index Terms*—tomato recognition; harvesting robot; K-means clustering; mathematical morphology**

I. INTRODUCTION

Agriculture product harvesting in greenhouses is a time consuming task for farmers. Many researches have been done to assist farmers' harvesting tasks. The development of robot system that enables harvesting autonomously has thus received considerable attention in the last decades. The use of robots to pick tree fruit was first proposed by Schertz and Brown in 1968 [1]. Kondo et al. [2], Hayashi and Sakaue [3] and Arima and Kondo [4] reported research prototypes of picking robots for tomatoes and cucumbers. The recently research on autonomous robots picking system have been addressed by Van Henten et al. [5]. The difficulties of picking robots system is the line-of-sight approach to fruit picking. that is, (1) to visually locate the fruit with an optical sensor; (2) to guide the fruit detachment device along the line of sight to the fruit; (3) to actuate the device when the fruit is contacted.

Our tomato picking robot is designed to pick up the ripe tomato in the green house without human intervention. A three dimensional movement robotic arm with a hand system is used as a mechanical system. The hand system is composed a scissor, a CCD camera and a laser sensor. To pick up the ripe tomato without human intervention the ripe tomato should be recognized and the accurate three dimensional coordinate of ripe tomato are required. With the accurate three dimensional coordinate and the size of tomato the robotic arm can be controlled to pick up the ripe tomato. In this paper, works are focus on the tomato recognition and extraction.

Image segmentation technology as an important object recognition method is uses to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze [6]. The result of image segmentation is a set of regions that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region are similar with respect to some characteristic or computed property, such as color, intensity, or texture [7]. Adjacent regions are significantly different with respect to the same characteristic. In this work, color feature is employed for ripe tomato segmentation. The ripe tomato are segmented by K-means clustering using the L*a*b* color space. To extract single integrity ripe tomato, mathematical morphology method is used to denoise and handle the conditions of tomato light overlapping and shelter.

The paper is organized as follows. Section 2 presents the proposed ripe tomato segmentation algorithm. And Section 3 describes the approach of single integrity ripe tomato extraction. Computer simulation and results analysis are presented in Section 4. Concluding remarks are given in Section 5.

II. RIPEN TOMATO SEGMENTATION

Several general-purpose algorithms and techniques have been developed for image segmentation, such as the clustering methods, the Histogram-based methods, the edge detection methods, the level set methods, the watershed transformation, model based segmentation, and neural network segmentation [6], [7], [8], [9]. But there is no general solution to the image segmentation problem, these techniques often have to be combined with domain knowledge in order to effectively solve an image segmentation problem for a problem domain.

In the practical applications of ripe tomato segmentation, dozens of tomato's images are taken in the green house. The typical image with ripe tomato shows as Fig. 1. From this image we can figure out that there are three main color sets regardless the lightness. The ripe tomato color is red; the unripe tomato, leaves and stalks color is green and the background color of green house is gray. A $L^*a^*b^*$ color space is a color-opponent space with dimension L for lightness, a^* and b^* for the color-opponent dimensions, based on nonlinearly-compressed CIE XYZ color space coordinates [10]. K-means clustering algorithm can be employed to cluster the tomato image to three different classifications based on the a^* and b^* color component which regardless with lightness. The classification including the red color can be considered as the candidate ripe tomato. Fig. 2 is the flow diagram of ripe tomato

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

Fig. 1. A typical image of ripe tomatoes.

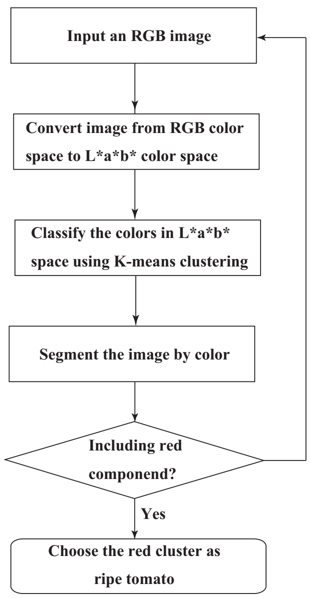

Fig. 2. Flow diagram of ripe tomato segmentation.

segmentation.

A. $L^*a^*b^*$ Color Space

Lab color space is a color-opponent space with dimension L for lightness and a and b for the color-opponent dimensions, based on nonlinearly-compressed CIE XYZ color space coordinates [10]. Unlike the RGB and CMYK color models, $L^*a^*b^*$ color is designed to approximate human vision. It aspires to perceptual uniformity, and its L component closely matches human perception of lightness. It can thus be used to make accurate color balance corrections by modifying output curves in the a and b components, or to adjust the lightness contrast using the L component [10] [11]. There are no simple formulas for conversion between RGB and $L^*a^*b^*$, because the RGB is a device dependent. The RGB first needs to be transformed to a specific absolute color space, such as sRGB or Adobe RGB. This adjustment will be device dependent, but the resulting data from the transform will be device independent, allowing data to be transformed to the CIE 1931 color space and then transformed into $L^*a^*b^*$.

Step1. The Transformation from sRGB color space to CIE XYZ color space Assuming the sRGB component values Rsrgb, Gsrgb, Bsrgb are in the range 0 to 1. (A range of 0 to 255 can simply be divided by 255). The formulas transform sRGB color space to CIE XYZ color space can be expressed as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} \begin{pmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{pmatrix}. \quad (1)$$

Step 2. The Transformation from CIE XYZ color space to CIE $L^*a^*b^*$ color space The formulas transform can be expressed as

$$\begin{aligned} L^* &= 116 f(Y/Y_n) - 16 \\ a^* &= 500[f(X/X_n) - f(Y/Y_n)]. \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)] \end{aligned} \quad (2)$$

where function $f(t)$ is defined as

$$f(t) = \begin{cases} t^{1/3}, & t > (6/29)^3, \\ \frac{1}{3}(\frac{29}{6})^2 t + \frac{4}{29}, & \text{otherwise}, \end{cases} \quad (3)$$

and X_n, Y_n and Z_n are the CIE XYZ tri-stimulus values of the reference white point.

The typical image with ripe tomato Fig. 1 can be transform to $L^*a^*b^*$ color space using above formula. Fig. 3 (a) is the L component of tomato image Fig. 3 (b) is the a component of tomato image. Fig. 3 (c) is the b component of tomato image and Fig. 3 (d) is the a and b components of tomato image. From the Fig. 3 we can clearly figure out that regardless the lightness, there are three main color clusters. The red cluster can be considered as ripe tomato candidate area.

B. Color-based Segmentation Using K-means Clustering

1) K-means Clustering: The k-means algorithm is an algorithm to cluster n objects based on attributes into k partitions, $k < n$ [12]. It is similar to the expectation-maximization algorithm for mixtures of Gaussians in that they both attempt to find the centers of natural clusters in the data. It assumes that the object attributes form a vector space. The objective it tries to achieve is to minimize total intra-cluster variance, or, the squared error function,

$$E = \sum_{i=1}^{k} \sum_{x_j \in S_i} (x_j - \mu_i)^2, \quad (4)$$

where there are k clusters $S_i, i = 1, 2, \cdots, k$ and μ_i is the centroid or mean point of all the points $x_j \in S_j$.

The steps of the proposed algorithm are as follows.

- Choose the number of clusters, k.

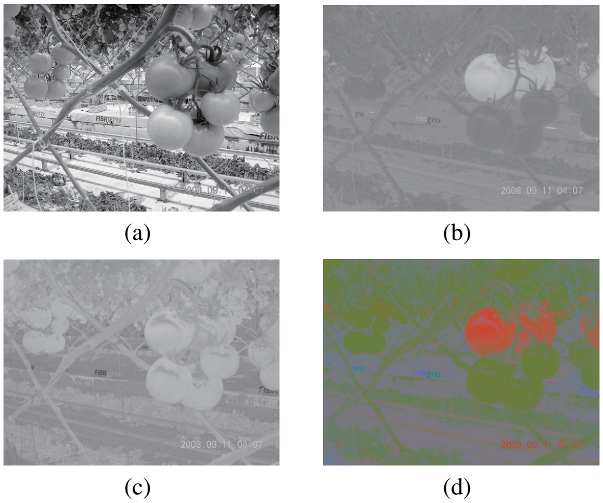

Fig. 3. The different component image under $L^*a^*b^*$ color space. (a) The L component image; (b) The a component image; (c)The b component image; (d) The ab components image.

- Randomly generate $k$clusters and determine the cluster centers, or directly generate k random points as cluster centers.
- Assign each point to the nearest cluster center.
- Recompute the new cluster centers.
- Repeat the two previous steps until some convergence criterion is met.

A drawback of the k-means algorithm is that the number of clusters k is an input parameter. An inappropriate choice of k may yield poor results. The algorithm also assumes that the variance is an appropriate measure of cluster scatter.

2) Proposed ripe tomato segmentation algorithm: Clustering is a way to separate groups of objects from complex background. K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. From the analysis in section 2.1, we know the typical image with ripe tomato can be separate to three groups by the a^* and b^* values. So three clusters to be partitioned and Euclidean distance metric is used as to quantify how close two objects are to each other. The proposed segmentation algorithm flow diagram show as Fig. 4.

For every tomato image inputed, k-means clustering algorithm returns an index corresponding to a cluster. The cluster center output from k-means clustering algorithm will be used later in this work. Label every pixel in the image with its cluster index. The labeled cluster index image shown as Fig. 5 (a). Using pixel-labels, objects can be separated by color, which will result in three images as Fig. 5 (b), Fig. 5 (c), Fig. 5 (d).

Find the image that contains the red component as the area of ripe tomato. Notice that there are dark and light red objects

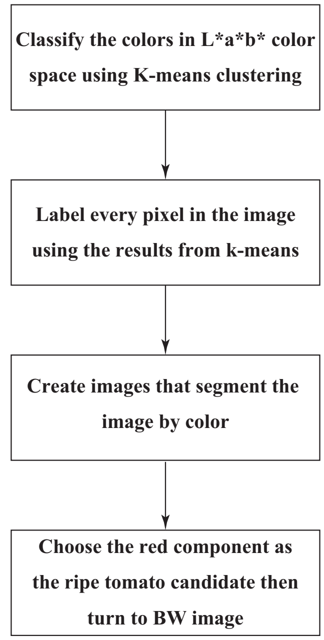

Fig. 4. Flow diagram of ripe tomato segmentation.

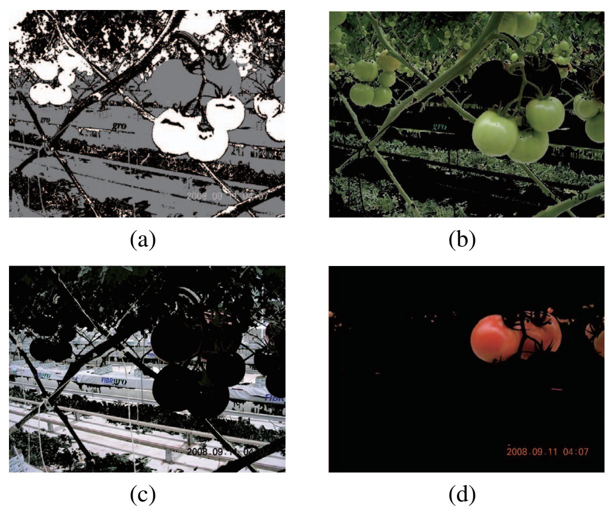

Fig. 5. The segmentation results based on proposed approach. (a) The labeled cluster index image; (b) The objects in cluster 1; (c) The objects in cluster 2; (d) The objects in cluster 3.

in this image. Recall that the L^* layer contains the brightness values of each color. The cluster-center contains the mean a^* and b^* value for each cluster. Hundreds ripe tomato color sample have been researched. Regardless the brightness values, the ripe tomato's a^* value is included in set (196, 146), b^* value is included in set (188, 132). Take the cluster which cluster-center value included in the statistical ripe tomato color set as the candidate ripe tomato area. Then turn the cluster image with ripe tomato into the binary image, show as Fig. 6.

Fig. 6. The binary image of a segmented ripe tomato.

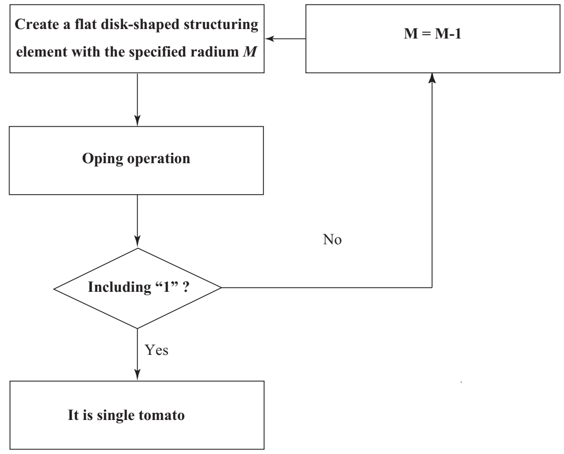

Fig. 7. Flow diagram of single integrity ripe tomato extraction.

III. The extraction of single integrity ripe tomato

A. Mathematical morphology

Mathematical morphology (MM) is a theory and technique for the analysis and processing of geometrical structures, based on set theory, lattice theory, topology, and random functions [13]. MM is most commonly applied to digital images, but it can be employed as well on graphs, surface meshes, solids, and many other spatial structures [14]. The basic operations are shift-invariant operators strongly related to Minkowski addition and subtraction. Let E be a Euclidean space or an integer grid, and A be a binary image in E.

- **Erosion**
  The erosion of the binary image A by the structuring element B is defined by:
  $$A \ominus B = z \in E \mid B_z \subseteq A, \tag{5}$$
  where B_z is the translation of B by the vector z. When the structuring element B has a center, and this center is located on the origin of E, then the erosion of A by B can be understood as the locus of points reached by the center of B when Bmoves inside A.
- **Dilation**
  The dilation of A by the structuring element B is defined by
  $$A \oplus B = \bigcup_{B} A_b. \tag{6}$$
  If B has a center on the origin, as before, then the dilation of A by B can be understood as the locus of the points covered by B when the center of B moves inside A.
- **Opening** The opening of A by B is obtained by the erosion of A by B, followed by dilation of the resulting image by B can be denotes as
  $$A \circ B = (A \ominus B) \oplus B. \tag{7}$$
  The opening is also given by $A \circ B = \bigcup_{B_x \subseteq A} B_x$, which means that it is the locus of translations of the structuring element B inside the image A.
- **Closing**
  The closing of A by B is obtained by the dilation of A by B, followed by erosion of the resulting structure by B:
  $$A \bullet B = (A \oplus B) \ominus B. \tag{8}$$
  The above means that the closing is the complement of the locus of translations of the symmetric of the structuring element outside the image A.

B. Proposed tomato extraction method

Using proposed ripe tomato segmentation algorithm, the results of the input tomato image are a group of tomatoes. But tomato picking robot is supposed to pick up the ripe tomato one by one. Extract the single integrity ripe tomato and get the tomato's center coordinate in the image is the key technology for tomato picking robot. Mathematical morphology method is used to separate the adhered ripe tomatoes and the biggest tomato area will be extracted as the fist pick up tomato. The algorithm flow is shown as Fig. 7.

Because of the tomato's shape is circle is the image. A flat disk-shaped structuring element with the specified radius M is created for the opening operation of the input binary image. The opening operation of input binary ripe tomato image A by created flat disk-shaped structuring element B is obtained by the erosion of A by B, followed by dilation of the resulting image by B. The erosion operation can separate the overlapping tomatoes by erode the edge of tomatoes. Then the dilation is used to dilate the tomato's size before the erosion operation. Parameter M is the radius of structuring element, which decide the erode area and dilate area size. Set a high value of M can change whole binary image's value to '0' . A property value of M can be used to get the biggest connected region. The iterate algorithm is proposed in Fig. 7 to get the property value of M. In this strategy the biggest single ripe tomato area is given. The results are shown as Fig. 8.

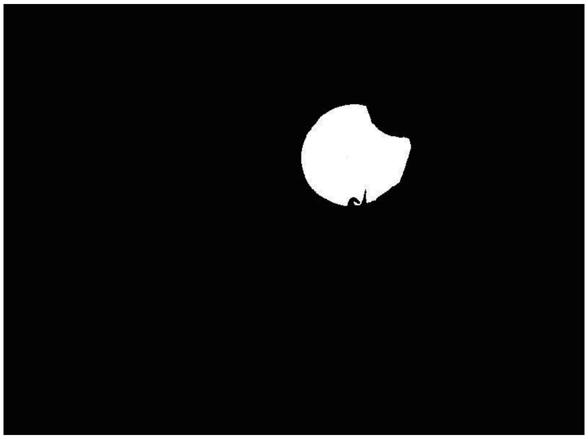

Fig. 8. Flow diagram of single integrity ripe tomato extraction.

IV. Experimental results

To illustrate the effectiveness of proposed method for the ripe tomato extraction. Hundreds tomato images are taken in the green house to test the tomato extraction algorithm. All the testify experiment are done using the MATLAB 7.0 and XP operation system in the laptop with 1GB RAM and 1.86G HZ Intel CPU. In order to get the average compute consume time, all the input images size are adjust to 800*600. In this experiment twenty tomato images are random chose from the images database. The algorithm average calculate time is 20.78100s. successful extract 19 images, failed one image. Some typical challenge extraction result shown in Fig. 9 and Fig. 10. and the failure sample shown in Fig. 11.

(a) (b) (c)

Fig. 9. The tomato image with Overlapping. (a) The input image; (b) The segmented tomato; (c) The extracted tomato.

(a) (b) (c)

Fig. 10. The tomato image with complex background. (a) The input image; (b) The segmented tomato; (c) The extracted tomato.

(a) (b) (c)

Fig. 11. The failed extraction example. (a) The input image; (b) The segmented tomato; (c) The extracted tomato.

V. Conclusion and discussion

In this work, the ripe tomato extraction algorithm are proposed based on the color clustering and mathematical morphology method. The experimental results data proved the proposed approach have an high performance in ripe tomato extraction. there are only one failed extraction from the twenty testify image samples. The conclusion can be drawn from the extract results of Fig. 9 and Fig. 10. Even in the lighter overlapping and complex background the proposed algorithm still have a very strong robustness.

There are also have some drawback of this proposed approach for ripe tomato extraction. such as, the failed sample shown as Fig. 11. The ripe tomato is just take litter part of the image it result the wrong cluster segmentation for the ripe tomato. The ripe tomato detective should be researched to make sure there always have the proper ripe tomato in the input images. Another import issue is the calculate consume problem, although the high performance computer can help to reduce the calculate time. But optimize the extraction approach is still the important work for the future research.

VI. Acknowledgment

This work is supported by Ontario Centres of Excellence.

References

[1] C. E. Schertz and G. K. Brown, “Basic considerations in mechanizing citrus harvest,” *Transactions of the ASAE*, vol. 11, pp. 343–348, 1968.

[2] M. Monta, N. Kondo and T. Fujiura, “Fruit harvesting robots in japan,” *Advances in Space Research*, vol. 18, pp. 181–184, 1996.

[3] O. Sakaue and S. Hayashi, “Basic considerations in mechanizing citrus harvest,” in *Annual International Meeting*. ASAE, pp. 963–967, 1996.

[4] N. Kondo and S. Arima, “Cucumber harvesting robot and plant training system,” *Journal of Robotics and Mechatronics*, vol. 11, pp. 208–212, 1999.

[5] E. J. Van Henten, J. Hemming, B. A. j. Van Tuijl, J. G. Kornet and J. Bontsema, “Collision-free motion planning for a cucumber picking robot,” *Biosystems Engineering*, vol. 86, pp. 135–144, Oct. 2003.

[6] D. L. Pham, C. Xu and J. L. Prince, “Current methods in medical image segmentation,” *Annual Review of Biomedical Engineering*, vol. 2, pp. 315-337, 2000.

[7] J. Im, J. R. Jensen and J. A. Tullis, “Object-based change detection using correlation image analysis and image segmentation,” *International Journal of Remote Sensing*, vol. 29, pp. 399-423, 2008.

[8] O. L. Vovk, “A new approach to visual similar image colors extraction,” *Journal of Automation and Information Sciences*, vol. 6, pp. 100–105, 2006.

[9] M. Brejl and M. Sonka, “Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples,” *IEEE Transactions on Medical Imaging*, vol. 19, pp. 100–105, 2000.

[10] M. D. Fairchild, “Color and image appearance models,” *Color Appearance Models*, pp. 340-399, 2008.

[11] G. Hoffmann, "Formulas from 'RGB' to CIELab," *CIELab Color space*, pp. 4-9, 2008.
[12] A. M. Fatih and E. Okan, "The performance factors of clustering ensembles," in *Signal Processing, Communication and Applications Conference*. IEEE, pp. 1-4, Apr. 2008.
[13] E. Aptoula and S. Lefvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, pp. 2914–2929, 2007.
[14] L. Vincent, "Morphological gray scale reconstruction in image analysis: applications and efficient algorithms," *IEEE Transactions on Image Processing*, vol. 2, pp. 176–201, 2002.