Computers and Electronics in Agriculture 94 (2013) 29–37

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

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

An image segmentation method for apple sorting and grading using support vector machine and Otsu's method

Akira Mizushima [a,*], Renfu Lu [b]

[a] United States Department of Agriculture, Agricultural Research Service, 524 S. Shaw Lane, Room 207, Michigan State University, East Lansing, MI 48824, USA
[b] United States Department of Agriculture, Agricultural Research Service, 524 S. Shaw Lane, Room 224, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:
Received 18 September 2012
Received in revised form 6 February 2013
Accepted 17 February 2013

Keywords:
Segmentation
Support vector machine
Otsu's method
Automatic thresholding
Apple sorting and grading

ABSTRACT

Segmentation is the first step in image analysis to subdivide an image into meaningful regions. It directly affects the subsequent image analysis outcomes. This paper reports on the development of an automatic adjustable algorithm for segmentation of color images, using linear support vector machine (SVM) and Otsu's thresholding method, for apple sorting and grading. The method automatically adjusts the classification hyperplane calculated by using linear SVM and requires minimum training and time. It also avoids the problems caused by variations in the lighting condition and/or the color of the fruit. To evaluate the robustness and accuracy of the proposed segmentation method, tests were conducted for 300 'Delicious' apples using three training samples with different color characteristics (i.e., orange, stripe, and dark red) and their combination. The segmentation error varied from 3% to 25% for the fixed SVM, while the adjustable SVM achieved consistent and accurate results for each training set, with the segmentation error of less than 2%. The proposed method provides an effective and robust segmentation means for sorting and grading apples in a multi-channel color space, and it can be easily adapted for other imaging-based agricultural applications.

Published by Elsevier B.V.

1. Introduction

Today computer vision technology has been widely used for quality inspection of agricultural and food products, replacing traditional manual operations that are labor intensive, slow and prone to human error (Sun, 2008). Despite the significant advances over the past decade, new image processing algorithms are continuously being developed to improve the accuracy and efficiency of automatic inspection for agricultural and food products (Baranowski et al., 2012; ElMasry et al., 2012; Garrido-Novell et al., 2012).

Image segmentation, which separates the product region from background in the image, is one of the most important tasks in image processing since it is the first step in image analysis after the image capture to subdivide an image into meaningful regions. The segmentation result affects the subsequent image analysis. For instance, the estimation of product size (Moreda et al., 2009) and shape (Moreda et al., 2012) is directly affected by the segmentation result because these morphological features are usually obtained based on the product contour information.

A number of image segmentation methods have been developed in the past (Zheng and Sun, 2008). Among them, thresholding-based segmentation is the most popular for online applications due to its simplicity and fast processing speed. The simplest thresholding selection method is manual selection. Li et al. (2002) applied manual thresholding after subtracting the background image. Manual thresholding, combined with an edge detection technique, was used to remove the background from the grayscale image for grading pre-sliced hams (Valous et al., 2009). Multi-thresholding methods using red and saturation were performed to filter the background for apples (Zou et al., 2010). Currently, manual thresholding method is still being used for image segmentation because of its simplicity. However, it is not practical or convenient in real world applications, where the optimal threshold may vary if the lighting condition and the object color change. Consequently, a number of automatic thresholding selection methods have been developed to adapt the changes in real time.

Otsu's thresholding selection method (1979) is one of the most accurate and widely used methods for image segmentation (Sahoo et al., 1988). Otsu's method automatically finds the threshold using the histogram of a grayscale image, based on the idea of finding the threshold that maximizes the between-class variance $\sigma_B^2(t)$ (or minimizes the weighted within-class variance), which is expressed as follows:

$$\sigma_B^2(t) = \frac{[m_G P(t) - m(t)]^2}{P(t)[1 - P(t)]} \quad (1)$$

where m_G is the average intensity of the entire image, $m(t)$ is the cumulative mean up to level t, $P(t)$ is the cumulative sum of

* Corresponding author. Tel.: +1 517 353 7793.
E-mail address: Akira.Mizushima@ars.usda.gov (A. Mizushima).

0168-1699/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.compag.2013.02.009

probability assigned to object (background). The optimum threshold is the value t^*, that maximizes $\sigma_B^2(t)$ as follows:

$$t^* = \underset{0 \leqslant t \leqslant L-1}{\arg\max}\,\sigma_B^2(t) \tag{2}$$

This method works well if the histogram is bimodal and the image has uniform illumination. Many researchers used this method for a variety of color spaces, including the hue color space (Nashat et al., 2011), grayscale images (Yang et al., 2009), near-infrared (NIR) images (Lunadei et al., 2012), hyperspectral images (Garrido-Novell et al., 2012), and linear combinations of RGB colors such as excess green, excess red (Meyer and Neto, 2008) and color index of vegetation extraction (Zheng et al., 2009). Since all these color or multi-channel spaces are selected on a trial-and-error basis with visual inspection, which is time-consuming, there is no specific and effective procedure to find the best color spaces to obtain good segmentation results, Mery and Pedreschi (2005) sought the best linear combination of RGB components that maximizes the variance of a grayscale image using a numerical gradient method. But they did not report the segmentation time, which is very important in online machine vision applications.

Classification-based segmentation is another popular method, especially for RGB color images. It is a pixel-oriented technique, in which each pixel is classified as either object or background. Linear discriminant analysis (or Bayesian classification) is the most popular classification-based segmentation method for color imaging (Blasco et al., 2003; Chinchuluun et al., 2009). This method generates a linear separation (or classification) hyperplane in a multi-dimensional space (e.g. 3-D for RGB color) based on the statistics of each region (object or background) by selecting a few representative pixels of the object or background. This hyperplane generation task, called training, is usually performed off-line before the online image processing is executed. Recently, support vector machine (SVM) is becoming a popular method to calculate the separation hyperplane for agricultural applications (Mitra et al., 2004; Nashat et al., 2011). Unlike linear discriminant analysis, support vector machine only uses "difficult points" close to the decision boundary, called support vector, and maximize the margin between support vectors. Because of this criterion, SVM is robust for untrained data. Although the hyperplane calculated by SVM is considered the best hyperplane, the method has limitation in adjusting to variations in the lighting condition and the color of products because the hyperplane is fixed.

The objective of this research was therefore to develop an automatic adjustable algorithm for segmentation of color images for apple sorting and grading, using Otsu's method and SVM. The method takes advantage of the adjustability of Otsu's method and the effectiveness of selecting the optimal classification hyperplane by SVM. The proposed algorithm consists of three main steps: (1) the optimal linear separation hyperplane in the 3-D RGB space is calculated by using linear SVM; (2) a contrast enhanced grayscale image is calculated from the linear hyperplane, and (3) the optimal threshold around the fruit boundary is automatically estimated using Otsu's method.

2. Materials and methods

2.1. Color segmentation by linear support vector machine

Support vector machine is a supervised machine learning algorithm originally designed to solve the two-group classification problems by generating the optimal separation hyperplane in a multi-dimensional space (Vapnik, 1995). The basic idea of SVM is to find the optimal hyperplane to separate a dataset, since, in theory, there exist many hyperplanes. A hyperplane is such chosen that the margin between it and the nearest data points, termed support vectors, of both classes is maximized. Because of this concept, linear SVM, which finds the "linear" optimal separation surface, is considered the best "linear" classification or segmentation algorithm in terms of accuracy and robustness for the untrained data. Although nonlinear SVM can solve a more complicated decision boundary, we selected linear SVM because the former requires more computational time when there are many support vectors. In contrast, the computational time for linear SVM classification is irrelevant to the number of support vectors.

The soft margin hyperplane algorithm (Cortes and Vapnik, 1995) is applied for non-separable data. Given a training set of instance-label pairs $(\mathbf{x}_i, t_i)$, $i = 1, \ldots, l$ where $\mathbf{x}_i \in R^n$ is the training sample and $t_i \in \{-1, +1\}$ is the class label, the linear SVM requires to find the solution of the dual representation of the maximum margin problem:

$$L_D(\alpha) = \sum_i \alpha_i - \frac{1}{2}\sum_{i,j} \alpha_i \alpha_j t_i t_j \mathbf{x}_i \cdot \mathbf{x}_j \tag{3}$$

Subject to:

$$0 \leqslant \alpha_i \leqslant C, \tag{4}$$

$$\sum_i \alpha_i t_i = 0 \tag{5}$$

where α_i is the positive Lagrange multipliers and parameter C (>0) controls the trade-off between the training error and the margin; a larger C means assigning a higher penalty to errors. α_i can be obtained by solving the quadratic programming problem of Eq. (3).

The separation hyperplane is given by:

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b \tag{6}$$

$$\mathbf{w} = \sum_{i=1}^{N_S} \alpha_i t_i \mathbf{x}_i \tag{7}$$

$$b = t_i - \mathbf{w}^{\mathrm{T}}\mathbf{x}_i \tag{8}$$

where $\mathbf{w}$ is the surface normal to the hyperplane, $|b|/||\mathbf{w}||$ is the perpendicular distance from the hyperplane to the origin, $||\mathbf{w}||$ is the Euclidean norm of $\mathbf{w}$, and N_S is the number of support vectors. The new data point $\mathbf{x}$ is classified by evaluating the sign of Eq. (6).

When linear SVM is applied to the color RGB image segmentation, it calculates a classification hyperplane in the 3-D RGB space shown in Fig. 1. Eq. (6) is rewritten as a linear combination of red, green and blue as follows:

$$Z(x,y) = w_R R(x,y) + w_G G(x,y) + w_B B(x,y) + b \tag{9}$$

$$w_R = \sum_{i=1}^{N_S} \alpha_i t_i R_i,\ \ w_G = \sum_{i=1}^{N_S} \alpha_i t_i G_i,\ \ w_B = \sum_{i=1}^{N_S} \alpha_i t_i B_i \tag{10}$$

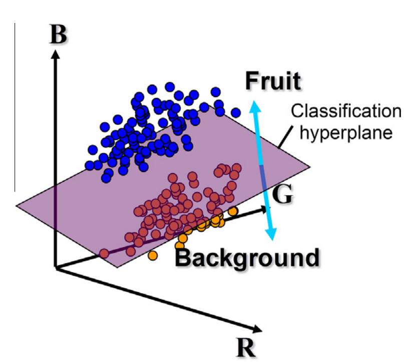

Fig. 1. Linear classification hyperplane in the 3-D RGB space.