Computers and Electronics in Agriculture 111 (2015) 92–102

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original paper

A key frame extraction method for processing greenhouse vegetables production monitoring video

Juncheng Ma, Xinxing Li, Haojie Wen, Zetian Fu, Lingxian Zhang *

College of Information and Electrical Engineering, China Agricultural University, Beijing, PR China
Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 3 July 2014
Received in revised form 15 November 2014
Accepted 10 December 2014

Keywords:
Greenhouse vegetables
Monitoring video
Key frames
Visual saliency
On line clustering
Mean pixels value

ABSTRACT

Research reported in this paper aims to improve the identification of greenhouse vegetable diseases based on the greenhouse monitoring video. It presents a method that combines the visual saliency and online clustering to extract the key frame from greenhouse vegetables monitoring video. Firstly X^2 histograms are used to measure the similarity of each frame to the first frame, which eliminates the meaningless frames and improve data processing efficiency and costs. Then, all frames will be converted to *HSV* color space and a saliency map of each frame is generated based on *H* component value and *S* component value. According to the saliency map, the salient region can be obtained. During the process of extracting the salient region, there is a possibility that the information of disease spots is lost. Therefore, morphological method would be utilized to restore the lost information. Finally, online clustering is performed to classify the salient regions into different clusters, and mean pixels value is used to select the key frames. The results indicate that this method can obtain information of entire leaf area of vegetables and extract the key frame effectively.

1. Introduction

Vegetable diseases are a common threat to farmers' agriculture production. Diseases will cause not only a decline on the quality of vegetables but also economic losses to greenhouse vegetables industry. Therefore, effective prevention, identification and treatment of disease are essential to ensure food safety and improve the yield and quality of the vegetable production (Pertot et al., 2012; Sankaran et al., 2010). The key point of effective prevention and treatment of disease is to identify the disease timely and accurately.

With the development of IOT (Internet of Things) technology, there is a trend that increasing number of agricultural sites has been equipped with video cameras monitoring the field situation and crop growth status, which allows the farmers to acquire visual information of the crop in greenhouse no matter where they are and when it is. Data collected by video cameras contains a lot of useful information, including diseases information, which is of great value in disease identification. However, the information has not been fully used. If these diseases information can be extracted from the monitoring videos for timely prevention and treatment, specific labors and investments on acquiring disease information can be saved; meanwhile it can significantly improve the vegetable production and reduce farmers' economic loss. Therefore, this paper extracted key frames from the greenhouse vegetables monitoring video to obtain the disease information.

A key frame is the most representative frame among a series of frames, which can describe the main content of a video shot briefly (Kim and Hwang, 2002). In this paper, a key frame is defined as a frame that contains completed disease spots information of vegetables foliage disease in the sequence of frames.

Currently, many key frame extraction algorithms have been proposed in the area of computer vision and image processing. Early key frame extraction methods are choosing key frames at predefined time point, which are not content-based and do not consider dynamics of the visual content and selected frames are often unstable (Walid and Ezzeddine, 2013). The average histogram algorithm and average pixels value algorithm are two of the conventional shot boundary based methods (Song et al., 2014). Although the computational costs of the two methods are low, they are not very effective at describing complex shots. Computing the frame differences based on some low level features is another early set of techniques. The frame will be chose as the key frame if its difference with the adjacent frames is more than a certain threshold (Ejaz et al., 2013, 2012; Jiang et al., 2009).

* Corresponding author at: College of Information and Electrical Engineering, China Agricultural University, PO Box 209, Qinghuadonglu 17, Haidian, Beijing 100083, PR China. Tel.: +86 10 627 36717.
E-mail addresses: zlx131@163.com, forever.mjc.cau@gmail.com (L. Zhang).

http://dx.doi.org/10.1016/j.compag.2014.12.007
0168-1699/

Optical flow analysis proposed by Wolf (1996) is a motion based analysis method, which is now adopted to meet the needs of different applications. For instance, Mendi et al. (2013) presented an automatic video summarization technique by motion metrics estimated from two optical flow algorithms, each using two different key frame selection criteria. Wang et al. (2012) proposed a new key frame extraction method based on optical flow and direction information entropy according to the characteristics of flower growing process. The optical flow based analysis method is required to process a large amount of data, especially if the object is moving in high speed that needs to extract many frames. Moreover, the threshold can have significant influence on the accuracy of the algorithm.

Clustering based methods are widely used in extracting key frames; especially the color feature based clustering method. For instance, Song et al. (2014) proposed a method for video abstraction based on fast online clustering of the regions of interest, which uses the hue histogram as the description feature for clustering. Kuanar et al. (2013) proposed an automated method of video key frame extraction using dynamic Delaunay graph clustering, where color histogram in *HSV* color space and texture feature by edge histogram descriptor are used as description features. De Avila and Lopes, 2011 used *K*-means algorithm to extract key frames by the Hue component histogram. Furini et al. (2010) proposed STIll and MOving Video storyboard based on the extraction of a 256-dimension vector that represents the average 256 bin colors histogram in the *HSV* color space of the video segment. Except for color feature, edge information can be used as description feature, too. For instance, Chan et al. (2011) used the edge structure of frames to eliminate the redundant key frames in case that the color information of two key frames is too similar. Clustering based methods are threshold-depending and impose that the number of key frames is set a priori (Walid and Ezzeddine, 2013).

All the key frame extraction method may work well in their specific field, but they may not suitable for extracting key frames from greenhouse vegetables monitoring videos because of the special characteristics. In this study, a novel method that took the characteristics of greenhouse vegetables monitoring videos into full consideration was proposed to extract the key frames from greenhouse vegetables monitoring videos.

In the research, X^2 was employed to eliminate the meaningless frames; then, IG algorithm was modified to obtain the salient region of each frame according to the characteristic of the videos. Finally, online clustering and mean pixel value was used to extract key frames.

2. Methodologies

2.1. Data acquisition

Six greenhouse vegetables monitoring videos were obtained from greenhouse Nos. 1 and 2 of Beijing Tongzhou Agricultural Science and Technology Demonstration Park by WAPA BL-C7HA1080IL-T18 High-speed infrared camera with a 2-mega-pixel-image sensor (Fig. 1). All the videos, 480×360 size, are obtained by daylight from 8:00 to 17:00 on 28, March 2014. The frame rate is 29 frames per second. The vegetables in the video are Spinacia oleracea.

The collected videos were divided into two groups: experimental videos (2) and test videos (4). The information of the videos is shown in Table 1.

2.2. Characteristics analysis of greenhouse vegetables monitoring videos

The application context determines what types of the key frame extraction algorithms should be used. To begin with we will provide a brief analysis on the characteristics of greenhouse vegetables monitoring videos which are summarized below (Fu et al., 2013):

Fig. 1. The camera.

(1) Brightness is the main variable in the monitoring videos.
(2) Motion feature of the monitoring cameras is the combination of slow movement and quiescent condition.
(3) The object monitored by the video cameras is one single vegetable plant.

After the characteristics analysis of greenhouse vegetables monitoring videos, the features of frames are discussed in following sections. Taking 2 videos as examples, this section shows the features of frames. Firstly, videos are decomposed to frames, and then two frames are chosen randomly, which are shown in Fig. 2.

As shown in Fig. 2, the characteristic of the two frames is very obvious, which means that the foreground, namely the vegetables, is clearly distinguished from background by their color regardless of the complexity of the background. The color feature of disease spots is very different from the normal leaf, which is likely to be mistaken as background. In *HSV* color space, *H* and *S* component value is invariant to illumination (Kuanar et al., 2013). Therefore, frames are converted to *HSV* color space to weaken the influence of illumination on object segmentation (Chen et al., 2011; Hu et al., 2012; Fu et al., 2013).

Assume that (r, g, b) is the value of a pixel in *RGB* color space, $r, g, b \in [0,255]$, $v' = \max(r, g, b)$ and define r', g', b':

$$r' = \frac{v' - r}{v' - \min(r,g,b)} \quad g' = \frac{v' - g}{v' - \min(r,g,b)} \quad b' = \frac{v' - b}{v' - \min(r,g,b)} \tag{1}$$

Define h':

$$h' = \begin{cases} (5 + b'), r = \max(r,g,b) \& g = \min(r,g,b) \\ (1 - g'), r = \max(r,g,b) \& g \neq \min(r,g,b) \\ (1 + r'), g = \max(r,g,b) \& b = \min(r,g,b) \\ (3 - b'), g = \max(r,g,b) \& b \neq \min(r,g,b) \\ (3 + g'), b = \max(r,g,b) \& r = \min(r,g,b) \\ (5 - r'), else \end{cases} \tag{2}$$