Computers and Electronics in Agriculture 151 (2018) 226–237

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Visual positioning technology of picking robots for dynamic litchi clusters with disturbance

Juntao Xiong[a], Zhiliang He[a], Rui Lin[a], Zhen Liu[a], Rongbin Bu[a], Zhengang Yang[a], Hongxing Peng[a], Xiangjun Zou[b,*]

[a] College of Mathematics and Informatics, South China Agricultural University, Guangzhou 510642, China
[b] College of Engineering, South China Agricultural University, Guangzhou 510642, China

ARTICLE INFO

Keywords:
Picking robot
Litchi
Dynamic target
Disturbance
Visual positioning

ABSTRACT

Visual localization by picking robots in natural environments is a difficult problem due to disturbance by interference factors. To solve this problem, research on visual positioning for dynamic litchi clusters with disturbance under natural environments was performed in this study. A visual system with two CCD cameras was used to acquire images of dynamic litchi clusters. The centroid of minimum bounding rectangle of litchi fruit was calculated, and the pendulum principle was used to calculate the oscillation angle of litchi clusters with three disturbance states: static, slight and large. For static or slight disturbance, the improved fuzzy C-means clustering method was used in image segmentation to obtain the litchi fruit and stem, and picking points were calculated using binocular visual stereo matching. The indoor experimental results show the maximum depth error of the picking point visual positioning for litchi clusters with static or slight disturbance is 5.08 cm, and the minimum depth error is 1.96 cm. The orchard visual positioning test results show the maximum depth error is 5.8 cm, and the minimum depth error is 0.4 cm. The results meet the picking demands of the picking robot end effector.

1. Introduction

Robot visual technologies have large technical differences for different robot functions. Industrial robots primarily perform visual detection and positioning of fixed targets or targets on a single production line in a specific factory environment (Veiga et al., 2013; Abu-Dakka et al., 2013), and agricultural robots primarily perform automatic operations for crops in natural environments. In terms of implementation technology, there are different visual technologies for different robots. In machine vision, there is a general theory for target detection and positioning using visual technology. For instance, tracking and positioning of dynamic targets in industry and agriculture is difficult.

In current agricultural production, research of visual positioning of picking robots primarily focuses on the identification and positioning of a static target in a natural environment or greenhouse (Zou et al., 2012), and dynamic targets with disturbance are studied less intensively. The randomness and uncertainty of dynamic targets makes visual identification and positioning difficult.

The current robot visual identification and positioning research of dynamic targets primarily involves agricultural product classification and detection of automatic production line (Milczarek et al., 2009; Kavdır and Guyer, 2008; Kurita et al., 2009; Momin et al., 2013), and little research exists for real-time identification and positioning of dynamic targets during robot picking operations. Visual positioning of oscillating apples was studied (Lü et al., 2012; Ning et al., 2015), and the fruit depth was measured using monocular vision in a lab environment with measurement error of less than 15 mm and a picking success rate of 84%. However, this research is based on the monocular visual positioning of a single fruit in a lab environment without in-depth analysis of laws of oscillation state and visual positioning. Due to differences in the fruit oscillation cycle, a monocular camera used to acquire images twice results in a time delay and positioning error. Meanwhile, this method is not suitable for visual positioning of fruit clusters with disturbance because the random structure of fruit clusters cannot be positioned by simply calculating the motion curve of each fruit but by identification and positioning of the stem of fruit clusters.

Disturbance is the random motion of an object by external force. Disturbance and static state are two natural states of fruits and vegetables in their growing environment. The primary reasons for disturbance are interfering factors in the natural environment, such as wind, mutual collision of branches and mechanical collision during robot picking. To achieve intelligent operation of a harvesting robot in a natural environment, interfering factors must be considered, and a coupling mechanism of its existing laws and accurate visual positioning

* Corresponding author.
E-mail address: xjzou1@163.com (X. Zou).

https://doi.org/10.1016/j.compag.2018.06.007
Received 11 October 2017; Received in revised form 21 May 2018; Accepted 2 June 2018

should be studied to establish an intelligent visual positioning decision-making mechanism and realize effective picking of fruits and vegetables (Ye et al., 2016). Studying visual identification and positioning of fruits and vegetables with disturbance in natural environments is important.

The current research of picking robots mainly focuses on identification and positioning of a single fruit, such as (Bulanon and Kataoka, 2010; Bac et al., 2014) oranges, apples, strawberries, and tomatoes, and less on fruit clusters, such as grapes, litchis, and longans (Luo et al., 2016; Wang et al., 2016). Identification and positioning of fruit clusters is more difficult than for a single fruit.

Litchi is a native fruit in south China, where a third of the world's litchis are produced. Designing a litchi picking robot to perform autonomous litchi harvesting can effectively reduce labor cost and improve the economic benefits of the litchi industry. Using litchi as a research object, by analyzing the disturbance status of litchi in natural environments, based on a picking robot end effector, this study uses a fruit centroid tracking method based on image processing to identify the disturbance state of litchi fruit, uses binocular stereo vision to determine the spatial location of the picking point of litchi clusters with disturbance, and provides technical support for dynamic target visual positioning of the picking robot.

2. Materials and methods

2.1. Robot visual system

The visual system includes two CCD cameras (model number DH-HV3100FC) with maximum resolution 2048 × 1536 pixels, pixel size 3.2 × 3.2 μm and a USB interface. The interior and exterior parameters of the cameras were determined using a calibration test, with a focal length of 8.2 mm and an optimal baseline distance of 120 mm between the two cameras.

2.2. Robot end effector structure

The robot has a self-designed end effector suitable for clamping and picking multiple types of fruit (Zou et al., 2016), shown in Fig. 1. The parameters of the end effector are a length of 60 mm and maximum opening width of 60 mm. The design allows for robot positioning and gripping of litchi fruits with a positioning error less than 60 mm along the x-axis and a position error less than 60 mm along the y-axis. The end effector has a certain error-tolerance, and fruits with disturbance can be picked successfully when the picking point is within the effective picking range of the end effector. For dynamic targets, the amplitude of disturbance can influence the picking success rate. Excessive disturbance will cause picking failure for dynamic targets beyond the picking scope of the end effector. When conducting visual positioning of a dynamic target with disturbance, the picking robot requires the target disturbance state to determine an accurate picking decision in real-time.

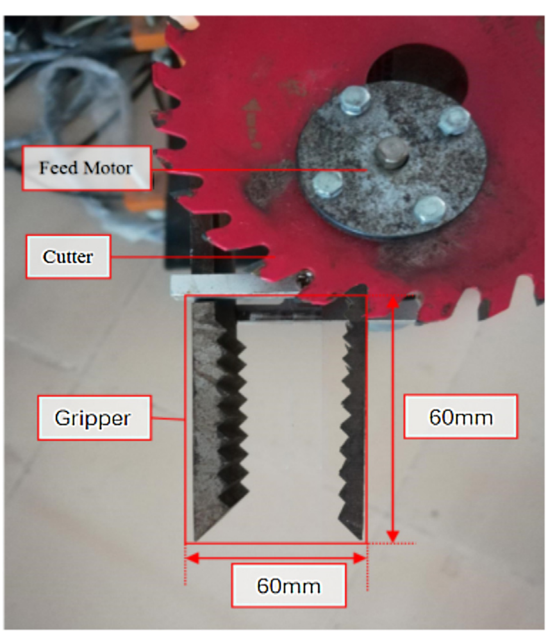

Fig. 1. End effector structure of litchi picking robot.

2.3. Structure of picking robot and control algorithm

Fig. 2 shows a working scene of the picking robot. In practice, the robot moves along the middle of the road and records images of the litchi trees on both sides of the road. The working process is as follows: According to the effective working distance of the manipulator, the highest and lowest point of the manipulator while obtaining images is determined. Next, the manipulator is controlled to move from top to bottom combining the camera angle. Images are obtained every other pre-set interval. Visual calculations are made for every image obtained. After that step, it is decided whether to pick up or give up according to the calculation result. After completing the work at the current height, the robot moves another interval to the next location and records another image. When the manipulator reaches the pre-set lowest point, it moves along the path and is repositioned to the highest point. The robot continues obtaining images.

In the actual work process, the picking robot can only pick one cluster at a time. Therefore, a visual location is made for only one cluster at a time. The remaining clusters are processed and images are obtained for another picking process. If there is more than one litchi cluster in the image, all litchi clusters are recognized. Next, the largest cluster is reserved in the image for visual location and the picking process.

During picking, the visual system is responsible for obtaining image coordinates of picking points and transforming them to world coordinates for the motion control system, in which the transform matrix of the manipulator motion is calculated according to the input world coordinates. Finally, the position of the end effector is controlled according to the transform matrix.

The world coordinates $W \triangleq [X_w, Y_w, Z_w, 1]^T$ are transformed to image coordinates $C \triangleq [X_c, Y_c, 1]^T$ with the transform matrix ${}^c_wT$ in formula (1). ${}^c_wT$ is a 4×3 matrix, the parameters of which are solved using Zhang's Calibration Method (Zhang, 2000). The matrix ${}^c_wT$ is used to locate the picking point by calculating the world coordinate with binocular vision.

$$C = {}^c_wT \cdot W \tag{1}$$

In the motion control system, the movement of the picking robot can be expressed as the relative positions of different connecting rods. The relative position of the *i*-th connecting rod and the *i*-1-th rod can be controlled with a 4 × 4 relative pose matrix ${}^{i-1}_{i}T$ of the neighboring rods.

As shown in Fig. 2, the picking robot used in this study has six joints and six rods. Therefore, the position of the end-effector $P \triangleq [X_p, Y_p, Z_p, 1]^T$ can be expressed in world coordinates as formula (2). According to formulas (1) and (2), the parameters in the position matrix can be solved using Paul's Kinematic Control Formulas (Paul et al., 1981).

$$W = {}^0_1T \cdot {}^1_2T \cdot {}^2_3T \cdot {}^3_4T \cdot {}^4_5T \cdot {}^5_6T \cdot P \tag{2}$$

After obtaining the position of the end-effector in world coordinates and the relationship between image coordinates and world coordinates, the relationship between the visual system and motion control system can be determined, as shown in Fig. 3.