BIOSYSTEMS ENGINEERING 145 (2016) 39–51

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Localisation of litchi in an unstructured environment using binocular stereo vision

Chenglin Wang [a], *Xiangjun Zou* [a,*], *Yunchao Tang* [b], *Lufeng Luo* [a], *Wenxian Feng* [b]

[a] *College of Engineering, South China Agricultural University, Guangzhou 510642, China*

[b] *School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China*

ARTICLE INFO

Article history:
Received 2 October 2015
Received in revised form
17 January 2016
Accepted 9 February 2016
Published online xxx

Key words:
Binocular stereo vision
Wavelet transform
Litchi recognition
Stereo matching

The major constraints for a litchi harvesting robot were to recognise and locate litchi in an unstructured environment with varying illumination and random occlusion. A rapid and reliable method based on binocular stereo vision was developed with the aim of effectively recognising and locating litchi in the natural environment. The method involved the application of wavelet transform to a pair acquired images of litchi to normalise illumination of an object surface. A litchi recognition algorithm based on *K*-means clustering was presented to separate litchi from leaves, branches and background. A matching algorithm to locate litchi based on a label template was discussed. Litchis with a similar label template were matched according to the preset threshold by traversing a litchi label template of a left image in a right image to find optimal matching. The experimental results showed that the proposed recognition method could be robust against the influences of varying illumination and precisely recognising litchi, the highest average recognition rate for unoccluded and partially occluded litchi was 98.8% and 97.5% respectively. From 100 pairs of tested images of unoccluded and partially occluded litchis 98% and 94% were successfully matched, respectively. Errors had no significant difference and they were less than 15 mm when the measuring distance was between 300 mm and 1600 mm under varying illumination and partially occluded conditions.

1. Introduction

Litchi is a kind of much-favoured fruit. The average annual production of litchi in the world is more than 200 million tonnes, about 80% of which is produced from China. However, with ageing population of China, problems of labour shortages and high labour-costs have been serious and they may even affect the smooth completion of litchi harvesting in the near future. In order to solve this problem, the research and development of a litchi harvesting robot is essential. Many researchers in other fields have developed fruit harvesting robots. Monta, Kondo, and Shibano (1995) developed an agricultural robot using four end-effectors for grape harvesting, berry thinning, spraying and bagging. The experimental

* *Corresponding author*. College of Engineering, South China Agricultural University, Wushan Road, Guangzhou 510642, China. Tel.: +86 020 85280066.

E-mail addresses: wangchenglin055@163.com (C. Wang), xjzou1@163.com (X. Zou).

http://dx.doi.org/10.1016/j.biosystemseng.2016.02.004

Nomenclature

PC	personal computer
CCD	charge coupled device
DOF	degree of freedom
R, G, B	red, green, and blue colour value
$C(x,y)$	R colour channel of original litchi colour image
x,y	the coordinate of the pixel point
k,l	independent variable that are both integers
g,h	high-pass filter and low-pass filter
c_j	the original image
R	the radius
d^D_{j+1}, c_{j+1}	diagonal high-frequency and low-frequency
x_0,y_0	coordinates of the centre of the circle
d^V_{j+1},d^H_{j+1}	vertical high-frequency and horizontal high-frequency
$P(x_0,y_0,R)$	an accumulator matrix
$P_{max}(x_0,y_0,R)$	the maximum accumulator matrix
A	the ratio of area
T	threshold for measuring partially occluded region size
b	the value of baseline
O (x_1,y_1,z_1)	world coordinates of litchi feature point
u_0, v_0, a_x, a_v	camera internal parameters
a,b	two portions that litchi C was segmented into
d	the value of $u_1 - u_2$, namely disparity
NCC	the normalised cross-correlation
$M \times N$	the size of the template
$I_1(u+i, v+j)$	the grey value of point $(u+i, v+j)$
Sig.	difference significance test value
$\overline{I_1}(u,v)$	the average value of grey values of label template
$\overline{I_2}(u-d,v)$	the average value of grey values of label template translation
z_1	distance measurement value between the litchi and cameras
(u,v)	the diagonal intersection point coordinates of litchi label template
$(u+i-d, v+j)$	the translation result of $(u+i, v+j)$ in right image

results showed that it could work efficiently. An orange picking robot constructed by Muscato and Prestifilippo (2005) involved a totally autonomous robot for fruit picking and handling crates. The picking time for the robotic fruit picker was at 8.7 per orange. Baeten, Donné, Boedrij, Beckers, and Claesen (2008) proposed an autonomous fruit picking machine (AFPM) for robotic apple harvesting. Its necessary components were assembled. It demonstrated the feasibility and functionality of the AFPM so that the picking cycle period was an average of 9. A litchi harvester designed by Liu, Zeng, and Ke (2011) was able to speed up litchi picking in the long term and picking by litchi harvester was significantly faster than hand picking tall litchi trees. An apple harvesting robot (De-An, Jidong, Wei, Ying, & Yu, 2011) consisted of a manipulator, end-effector and image-based vision servo control system autonomously performed its harvesting task using a vision-based module and picked the apples by controlling a manipulator with 5 DOF structure. The success rate of apple harvesting was 77%, and the average harvesting time was approximately 15 per apple. A strawberry harvesting robot system (Feng, Liu, Wang, Zeng, & Ren, 2012; Feng, Zheng, Qiu, Jiang, & Guo, 2012) based on machine vision, sonar technology and an independent navigation system was found to harvest strawberry from both sides and distinguish and locate fruit by a vision system. A successful harvesting rate of 86% was achieved, every successful harvesting operation on average took 31.3, and the average error for fruit localisation was <4.6 mm. Robotic fruit harvesting was reviewed by Bac and van Henten (2014).

Recognition and localisation of fruit is an important basic procedure for developing a fruit harvesting robot (Mehta & Burks, 2014). In most reported studies, the camera is a key element of the vision system of the robot playing a crucial role in producing images for the robot to recognise and locate fruits. Some studies concerning the ability of the vision system of the robot to recognise fruits have been reported. In Slaughter and Harrell (1987), over 75% oranges were detected through enhancing the contrast between orange fruits and others by a colour camera with a filter of 675 nm wavelength, but some misclassified non-fruit part were presented. Grey level fruit images of peach and apple were obtained by a single CCD camera with filters, and up to 92% fruits were detected by Sites Peter and Delwiche Michael (1988). Zhao, Tow, and Katupitiya (2005) analysed a single colour image to segment apples using both colour and texture, the correct rate was about 90%. Although these methods were able to recognise fruits, they did not have sufficient ability to adapt to the fruit harvesting because the analysis was still based on two-dimensional images which could not determine the distance from the camera to the image scene.

The binocular stereo vision is a more favourable method to carry out the three-dimensional localisation of fruit. It includes two cameras using different views to capture two images, thus the distance between the fruit and the camera can be calculated by triangulation. In reported studies, images have been captured by two parallel cameras or cameras at specified angles, thus the distance from the camera to the image scene could be calculated by combining the left and right images. An early stereo vision system developed for an apple harvesting robot involved the use of two cameras positioned at an angle with 40% of visible apples recognised in the experiment (Kassay, 1992). A similar scheme (Grasso & Recce, 1996) was implemented so that oranges could be recognised by a stereo matching algorithm using two cameras to capture images from different view of angles. Another binocular stereo vision system for robotic apples harvesting was developed using two cameras placed in parallel (Teruo, Shuhuai, & Hiroshi, 2002). The rate of fruit discrimination was about 90% for red fruits and 65%–70% for yellow-green apples. This method was also used in several studies for recognition and localisation of fruits. Xiang, Jiang, and Ying (2014) proposed a binocular vision system for recognising clustered tomatoes. The depth images of clustered tomatoes were acquired through the two colour cameras placed in parallel and the recognition accuracy rate of clustered tomatoes was 87.9% when the leaf or branch occlusion rate was