



# Integration of stereo vision system calibration and kinematic calibration for an autonomous kiwifruit harvesting system

Chi Kit Au<sup>1</sup> · Shen Hin Lim<sup>1</sup> · Mike Duke<sup>1</sup> · Ye Chow Kuang<sup>1</sup> · Michael Redstall<sup>1</sup> · Canaan Ting<sup>1</sup>

Received: 28 May 2022 / Accepted: 29 September 2022  
© The Author(s) 2022

## Abstract

Stereo vision system and manipulator are two major components of an autonomous fruit harvesting system. In order to raise the fruit-harvesting rate, stereo vision system calibration and kinematic calibration are two significant processes to improve the positional accuracy of the system. This article reviews the mathematics of these two calibration processes and presents an integrated approach for acquiring calibration data and calibrating both components of an autonomous kiwifruit harvesting system. The calibrated harvesting system yields good positional accuracy in the laboratory tests, especially in harvesting individual kiwifruit. However, the performance is not in line with the outcomes in the orchard field tests due to the cluster growing style of kiwifruit. In the orchard test, the calibrations reduce the fruit drop rate but it does not impressively raise the fruit harvesting rate. Most of the fruit in the clusters remain in the canopy due to the invisibility of the stereo vision system. After analyzing the existing stereo vision system, a future visual sensing system research direction for an autonomous fruit harvesting system is justified.

**Keywords** Stereo vision system calibration · Kinematic calibration · Robotics · Harvesting

## 1 Introduction

The growing New Zealand kiwifruit industry has achieved export sales of \$2.599 billion NZD during the 2020/21 season (Zespri 2020) and the revenue is expected to be \$6 billion NZD by 2030 (NZKGI 2020). With this rapid growth, kiwifruit growers are experiencing challenges associated with labour shortages and rising labour costs. Labour shortages have become apparent during the 2018 harvesting season, as declared by the New Zealand Ministry of Social Development of an official seasonal labour shortage of 1200 people for six weeks (Msd 2018). This shortage is likely to worsen with a projected 7000 more seasonal workers required for the kiwifruit industry by 2027 (NZKGI 2018). The second problem is rising labour costs. Pruning, thinning, and picking wages make up approximately half of the total orchard working expenses based on a 2011/12 orchard model. The New Zealand minimum wage has significantly increased recently. The Government has largely increased

the minimum wage by 2021 (Statistics New Zealand 2018). Of the 2021 labour expenses, pruning wages are the greatest at \$9,700 NZD/ha, followed by picking wages at \$3,922 NZD/ha (Mpi 2021).

Robotics and automation in agriculture can play a key role in the long-term solution to both the labour shortage and rising labour cost. As a result, an autonomous kiwifruit harvesting system is developed in-house as shown in Fig. 1. The configuration parameters of this harvesting system can be found in references (Williams et al. 2019, 2020). The system consists of four identical modules. Each module has a robot and a stereo vision system. The robot has a manipulator for moving the gripper, which grabs the kiwifruit and a corrugated tube for delivering the kiwifruit to the container. Although each module has a stereo vision system, it is not just dedicated to the robot of that module. The information gathered by the stereo vision system in each module will be merged together by a scheduler, which assigns the fruit to be harvested to each robot. This approach can fully utilize the resources to complete the task with the minimum time.

This autonomous kiwifruit harvesting system has been employed for fruit harvesting in various kiwifruit orchards. The harvesting task by robot in an orchard is quite similar to the pick and place operation in the manufacturing

✉ Chi Kit Au  
ckau@waikato.ac.nz

<sup>1</sup> University of Waikato, Hamilton, New Zealand



**Fig. 1** The kiwifruit harvesting system with four robot modules

**Table 1** The categories for harvesting attempts in the previous video footages

Number of attempts	Harvested (%)	Dropped (%)	Unharvested (%)
1456	50.8	24.6	24.6

environment, but it is more challenging due to uncontrollable dynamic conditions in orchards. For instance, no two canopies are the same, the robot base is not fixed as it has to navigate the orchards with many obstacles to delicately grab and place a fruit.

One of the significant factors that affects the economic viability of the system is the successful harvesting rate of the total detached fruit. All the fruit left in the canopy and dropped on the ground are considered unsuccessful harvesting, and thus, dropped kiwifruit during harvesting have a significant impact on the orchard profit. Fruit drop study is commissioned to investigate the number and causes of unsuccessful harvesting. A total of 1456 harvesting attempts under various operation conditions in ten hours of video footage is watched and recorded manually. The attempts are categorized into picked, dropped and unharvested as listed in Table 1.

The main reason for fruit left in the canopy and dropping on the ground is position error between the gripper and the fruit location. This is a combination of inaccuracies due to fruit localization by the stereo vision system, kinematics of

the robot and the coordination between the vision system and the robot. Calibration is one of the approaches to remedy these inaccuracies.

This article presents an integrated approach to obtain the calibration data and perform these calibration processes for the autonomous kiwifruit harvesting system so that its positional errors are reduced. Field tests are carried out in the kiwifruit orchard. The performance of the calibrated harvesting system has shed light on future visual sensing system research for an autonomous fruit harvesting system.

## 2 Related work

Various robots are developed for harvesting different fruits such as tomatoes (Liu et al. 2015), strawberry (Xiong et al. 2020), apples (Silwal et al. 2017), kiwifruit (Mu et al. 2020) etc. These robots are demonstrated in the lab environment and they are prototypes for design concept proofs. Most of them are just one robot arm with a camera for sensing a target fruit. They are not fully autonomous to sense over hundreds of fruits in the canopy simultaneously and determine the harvesting order.

In order to develop an industrial scale autonomous harvesting system, harvesting speed and successful harvesting rate are two major factors to be considered. The system should have multiple robot arms to perform the harvesting task simultaneously so that the harvesting speed is

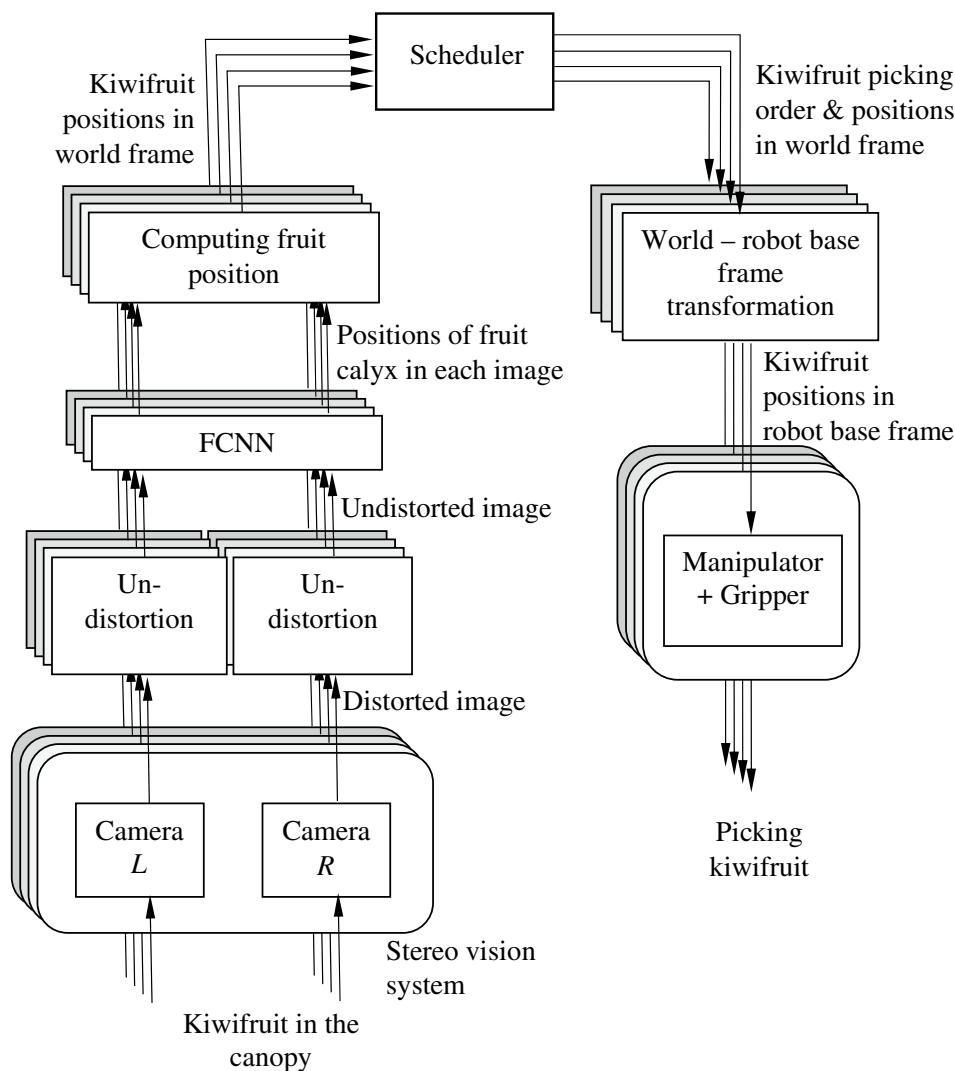
comparable with the manual harvesting speed. This implies that the canopy view captured by the sensing system such as a set of stereo vision systems must be as much as possible, since an autonomous robotic system for kiwifruit harvesting must first perceive and then correctly position the gripper precisely to perform a task. Furthermore, a scheduler is required for efficient fruit assignment to each robot. The information flow for harvesting kiwifruit with four modules is outlined in Fig. 2.

The harvesting robotic system is positioned under the canopy to capture the kiwifruit distribution. The canopy is imaged by both camera L and R of the stereo vision system in each module. Since the images obtained from the cameras are distorted due to the lens geometrics, the images will be undistorted before the kiwifruit are detected based on the calyx recognition by using a fully convolution neural network (FCNN). The positions of the detected kiwifruit are calculated *with respect to the world frame*. These positions from each module are fed into a scheduler so that all the fruit

positions are *merged*. Then the kiwifruit is assigned to each robot arm for harvesting based on the minimum distance between the fruit and the current gripper position. Once the kiwifruit picking order is scheduled, the corresponding fruit position with respect to the world frame is transformed to the robot base frame. For each fruit position, the joint parameters are computed by inverse kinematics. The arm attempts to drive the gripper to harvest the kiwifruit when it receives the joint parameters sequentially.

The high successful harvesting rate refers to positioning the gripper accurately to the target kiwifruit for detachment from the canopy. When there is a discrepancy between the gripper position and the actual kiwifruit position, the fruit may not be correctly grabbed. These errors in the system have a negative impact on the successful harvest rate which determine the economic viability of the robotic system. A common approach for mitigating the positional inaccuracies is through calibration so that positional difference between the commended position and the actual position is reduced.

Fig. 2 Information flow in the kiwifruit harvesting system with four modules



Stereo vision system calibration, kinematic calibration and world—robot base frame transformation estimation are the major calibrations for reducing the positional inaccuracies.

A stereo vision system consists of two cameras for sensing the depth information of a point. The computation is only related to the image locations in two cameras with their intrinsic and extrinsic parameters (Feng and Fang 2021). As a result, camera calibration is needed to obtain these parameters prior to the stereo vision system calibration. A pinhole camera model is used for modelling a camera (Brown 1971) which relates a 3D point with its image location on the image plane by a projection matrix. The pinhole camera model is not directly applicable when the camera lens causes a nonlinear distortion in the image. A simplified polynomial model (Tsai 1987; Zhang 2016) is used for image undistortion so that the computation efficiency can be raised. The parameters of this model have to be estimated together with the camera intrinsic and extrinsic parameters using Levenberg–Marquardt Algorithm (Zhang 2016; Moré 1987). The calibration process is equivalent to estimating the projection matrix using a set of 3D calibration points and their corresponding image locations. Various methods are employed to acquire the calibration points from a calibration tool (Tsai 1987; Zhang 2016).

Calibrating a stereo vision system refers to estimating the relative position and orientation of one camera with respect to the other camera. This can be determined from their orientations and positions obtained in the individual camera calibration procedures (Zhang et al. 2009; Cui et al. 2016). On the other hand, this information is also completely described by the epipolar geometry of the stereo vision system (Zhang 1998) which relates the image locations on the right and left image plane of a 3D point by a fundamental matrix. The orientation and position of one camera expressed in the other camera coordinate frame can be obtained directly from the fundamental matrix (Zhang 2016). Furthermore, stereo camera calibrations are also developed based on the neural network relating the 3D point with its image location (Memon and Khan 2010). It is not necessary to model the distortion separately in this approach, but the noise and instabilities are also included in the network.

The major error in the position and orientation of the end-effector is due to the robot component tolerances and clearances which are termed geometric variation. The error can be expressed in terms of the geometric variation through the kinematic model of a robot. Kinematic calibration is a process of obtaining the geometric variation of a robot so that the error of the end-effector position and orientation is minimized. The kinematic model of a robot depends upon its mechanical structure. For instance, it can be obtained by using Denavit–Hartenberg parameters (Rocha et al. 2011). On the other hand, the

non-geometric parameters such as the manipulator stiffness (Lightcap et al. 2008) and the compliance in the joints can also be included in the kinematic model (Nubiola and Bonev 2013). As a set of calibration points is needed for calibration, various methods are proposed to acquire these points. These include dial gauges (Xu and Mills 1999), laser trackers (Park et al. 2008) and 3-D camera (Motta et al. 2001) etc.

In summary, the modelling of various calibrations with the calibration data leads to a system of linear equations. In order to eliminate the noise influence, this system of linear equations is overdetermined. The research focuses on the autonomous robotic system accuracy including the coverage of the process and component models, computation efficiency of solving the system of equations and approaches of acquiring calibration data so that the error accumulation can be minimized.

### 3 Mathematical models for calibration

The mathematics for formalizing the calibration of the stereo vision system, kinematic and integration of both systems are briefed. Figure 3 depicts the configuration of a module which is part of the autonomous kiwifruit harvesting system. The robot has a manipulator and a gripper while the stereo vision system consists of two cameras  $R$  and  $L$ . There are four coordinate frames.  $\{F_w\}$  is the world frame,  $\{F_b\}$  is the robot base frame,  $\{F_r\}$  and  $\{F_l\}$  are the frame of camera  $R$  and  $L$  respectively. Furthermore, the frame  $\{F_r\}$  of camera  $R$  is also the stereo vision system frame. The relationships among these four frames  $\{F_w\}$ ,  $\{F_b\}$ ,  $\{F_r\}$  and  $\{F_l\}$  are:

$$\{F_l\} = {}_l^r T \cdot \{F_r\} \quad (1a)$$

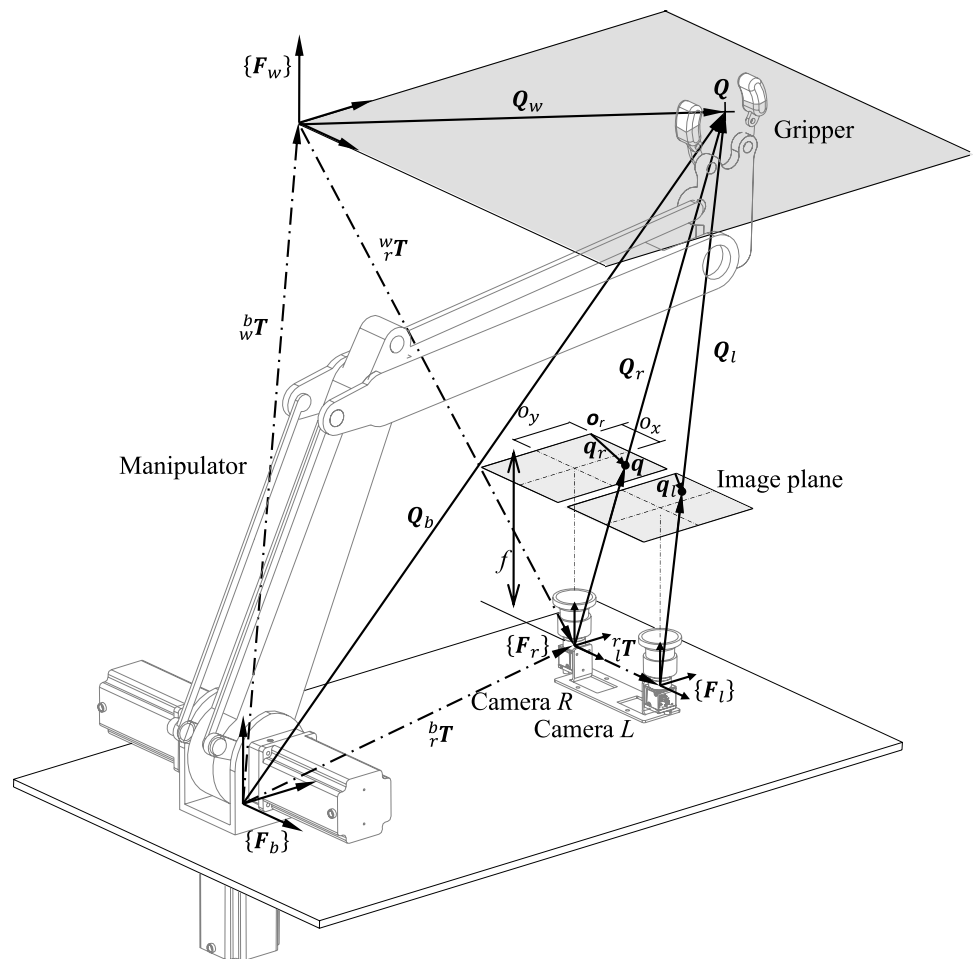
$$\{F_r\} = {}_r^b T \cdot \{F_b\} \quad (1b)$$

$$\{F_w\} = {}_w^b T \cdot \{F_b\} \quad (1c)$$

where  ${}^i T$  is a transformation matrix of frame  $\{F_i\}$  with respect to  $\{F_j\}$ .

A point  $Q$  located on the  $xy$  plane of world frame  $\{F_w\}$  has a position vector  $Q_w = [x_w \ y_w \ 0]^T$ . Its image  $q$  is on the image plane which is at the focal length  $f$  along the  $z$  axis of the camera frame  $\{F_r\}$ . It has a position vector  $q_r = [u_r \ v_r]^T$  and  $q_l = [u_l \ v_l]^T$  on the image plane of camera  $R$  and  $L$  with point  $o_r$  and  $o_l$  as the origin of the pixel frame respectively. The principal point, where the optical axis pierces the image plane, is at  $(o_x, o_y)$  from the origin  $o$ .

**Fig. 3** A module of the autonomous kiwifruit harvesting system



The stereo vision system calibration is split into two parts. The calibrations are performed on the cameras in the stereo vision system and the system itself.

### 3.1 Camera calibration

Camera calibration is to obtain the intrinsic and extrinsic parameters of the cameras. The intrinsic parameter matrix is

$$K = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

which consists of the pixel focal lengths  $f_x$  and  $f_y$  of the camera (focal length  $f$ ) in  $x$  and  $y$  direction of the image plane. The extrinsic parameters are the orientations and positions of the cameras.

Consider a point on the  $xy$  plane of the world frame  $\{F_w\}$  so that its  $z$  coordinate is zero. Camera  $R$  is at an orientation  ${}^rR = [{}^r r_1 \ {}^r r_2 \ {}^r r_3]$  and a position  ${}^r t$

expressed in the world frame  $F_w$  such that  ${}^rT = \begin{bmatrix} {}^wR & {}^w t \\ 0 & 1 \end{bmatrix}$ .

The point  $Q$  at  $(x_w, y_w, 0)$  in the world frame is related to its image  $q$  at  $(u, v)$  in the pixel frame on the image plane as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \tag{3}$$

where  $s$  is a scale factor and  $H = [h_1 \ h_2 \ h_3]$  is the homography projection matrix and  $h_i^T$  is the  $i$ -th row of homography  $H$ . The pixel coordinates  $(u, v)$  are un-distorted using Levenberg–Marquardt algorithm (Weng et al. 1992). The homography  $H$  is computed by stacking Eq. (3) for each calibration point  $Q_w^i$  ( $i = 1, 2, \dots, n$ ) from a single view yields a system of homogenous equations

$$A_h \cdot h = 0 \tag{4}$$

where  $A_h$  is a  $n \times 9$  coefficient matrix and  $h = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$ . The

homography  $H$  can be estimated by solving Eq. (4). On the other hand, the homography  $H$  can also be expressed as

$$H = K \cdot \begin{bmatrix} {}_r^w r_1 & {}_r^w r_2 & {}_r^w t \end{bmatrix} \quad (5)$$

The intrinsic parameter matrix  $K$  and extrinsic parameters  ${}_r^w R$  and  ${}_r^w t$  of a camera can be decomposed from homography  $H$  using a number of  $m$  ( $m \geq 2$ ) views. Each view  $j$  has a different homography  $H^j$  (and so do  $R^j$  and  $t^j$ ). However,  $K$  is the same for all views. The homography for each view,  $H^j = K \cdot \begin{bmatrix} {}_r^w r_1^j & {}_r^w r_2^j & {}_r^w t^j \end{bmatrix}$  provides two linear equations in the 6 entries of the matrix  $D = K^{-T} \cdot K^{-1} = [d_1 \ d_2 \ d_3]$  ( $\therefore$ Disymmetricmatrix).

Letting  $w_i^j = K \cdot {}_r^w r_i^j$ , the rotation constraints  $({}_r^w r_i^j)^T \cdot {}_r^w r_i^j = 1$  ( $\forall i = 1, 2$ ) and  $({}_r^w r_1^j)^T \cdot {}_r^w r_2^j = 0$  implies

$$({}_r^w r_1^j)^T \cdot D \cdot w_1^j - ({}_r^w r_2^j)^T \cdot D \cdot w_2^j = 0 \quad (6a)$$

$$({}_r^w r_1^j)^T \cdot D \cdot w_2^j = 0 \quad (6b)$$

Stacking 2  $m$  equations from  $m$  views, a linear system of equations is yielded

$$A_z \cdot d = 0 \quad (7)$$

where  $A_z$  is a  $2m \times 6$  coefficient matrix and  $d = [d_{11} d_{12} d_{22} d_{13} d_{23} d_{33}]^T$  (which are the components of  $D$ ). The intrinsic parameter matrix  $K$  can be obtained from  $D$  by Cholesky decomposition.

$$D = K^{-T} \cdot K^{-1} \quad (8)$$

The extrinsic parameters for view  $j$  can be computed using  $K$  and  $H^j = [h_1^j \ h_2^j \ h_3^j]$ ,

$${}_r^w R^j = [K^{-1} \cdot h_1^j \ K^{-1} \cdot h_2^j \ K^{-1} \cdot h_3^j \times K^{-1} \cdot h_2^j] \quad (9a)$$

$${}_r^w t^j = K^{-1} \cdot h_3^j \quad (9b)$$

### 3.2 Stereo vision system calibration

The orientation and position of a camera in the stereo vision system relative to the other is estimated during the *stereo vision system calibration* process. As shown in Fig. 3, frame  $\{F_r\}$  of camera  $R$  is the stereo vision system frame,  ${}_l^r R$  and  ${}_l^r t$  are the orientation and position of camera  $L$  expressed in the frame  $\{F_r\}$  of camera  $R$  such that  ${}_l^r T = \begin{bmatrix} {}_l^r R & {}_l^r t \\ 0 & 1 \end{bmatrix} \cdot K_r$  and

$K_l$  are the matrix  $K$  (Eq. (2)) of camera  $L$  and  $R$  respectively, then

$$[q_r^T \ 1] \cdot F \cdot \begin{bmatrix} q_r \\ 1 \end{bmatrix} = 0 \quad (10)$$

where  $F = (K_l^{-1})^T \cdot E \cdot K_r^{-1} = [f_1 \ f_2 \ f_3]$  is the fundamental matrix. By substituting  $n$  pairs corresponding image points in Eq. (10), the components of fundamental matrix  $F$  can be formulated a set of homogenous equations

$$A_f \cdot f = 0 \quad (11)$$

where  $A_f$  is a  $n \times 9$  coefficient matrix and  $f = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}$ . The

estimated fundamental matrix  $F$  are obtained by solving Eq. (11) which also gives the essential matrix  $E$ . Consequently, the orientation  ${}_l^r R$  and position  ${}_l^r t$  of camera  $L$  expressed in the frame  $\{F_r\}$  of camera  $R$  is calculated from  $E$  using single value decomposition.

$$E = (K_l)^T \cdot F \cdot K_r \quad (12a)$$

$$E = T_x \cdot {}_l^r R \quad (12b)$$

where  $T_x$  is the skew-symmetric matrix of  ${}_l^r t$ .

### 3.3 Kinematic calibration

The end-effector location  $Q_b$  with respect to the robot base frame  $\{F_b\}$  is expressed in terms of the robot geometric  $a$  and joint parameters  $\theta$ ,

$$Q_b = f(a, \theta) \quad (13)$$

The position error  $\Delta Q_b$  of the end-effector due to a set of identified geometric variation  $\Delta a$  is given by

$$\Delta Q_b = J \cdot \Delta a \quad (14)$$

where  $J$  is the Jacobian matrix of the kinematic model  $f$ . The kinematic calibration is to determine the robot geometric variation  $\Delta a$  so that the position error  $\Delta Q_b$  is eliminated.

Using a set of  $n$  end-effector locations at various corresponding joint parameters, Eq. (14) can be expressed in form of a set of homogeneous equations with geometric variation  $\Delta \bar{a} = [\Delta a_1 \ \dots \ \Delta a_m]^T$  (where  $m$  is the number of identified geometric variations)

$$\begin{bmatrix} J^1 \\ \vdots \\ J^n \end{bmatrix} \begin{bmatrix} \Delta a_1 \\ \vdots \\ \Delta a_m \end{bmatrix} = \begin{bmatrix} \Delta Q_b^1 \\ \vdots \\ \Delta Q_b^n \end{bmatrix} \text{ or } A_k \cdot \Delta \bar{a} = \Delta \bar{Q}_b \quad (15)$$

where  $A_k = \begin{bmatrix} J^1 \\ \vdots \\ J^n \end{bmatrix}$  and  $J^i$  is the Jacobian matrix of kinematic model  $f$  with  $i$ -th configuration;  $\Delta \bar{Q}_b = [\Delta Q_b^1 \dots \Delta Q_b^n]^T$ .

Hence, the geometric variation  $\Delta \bar{a}$  can be estimated by solving Eq. (15).

### 3.4 Integration calibration

The stereo vision system in the autonomous robotic system returns the coordinates of a target location with respect to its frame  $\{F_r\}$ . These coordinates are transformed into the world frame  $\{F_w\}$  by the transformation matrix  ${}_r^wT$  (obtained from the camera calibration) so that locations from various stereo vision systems are merged together in the world frame for harvesting scheduling. However, the robot control modules orientates and moves the end-effector to a target location expressed in the robot base frame  $\{F_b\}$ . The world-robot base frame transformation is a process that binds the domain of locations (expressed in the world frame) perceived by the stereo vision system with the domain of end-effector target locations in robot base frame. These two tasks need to be integrated by a transformation matrix  ${}_w^bT$ , which is estimated based on a set of  $n$  calibration points.

A set of  $n$  points  $Q^i (\forall i = 1, \dots, n)$  expressed in the world frame with position vector  $\{Q_w^i = [x_w^i \ y_w^i \ z_w^i]^T, \forall i = 1, \dots, n\}$  and the robot base frame with position vector  $\{Q_b^i = [x_b^i \ y_b^i \ z_b^i]^T, \forall i = 1, \dots, n\}$  are employed to establish the transformation matrix.

A rigid correspondence between these two sets of coordinate expression (i.e. point  $Q^i$  is always expressed as  $Q_b^i$  in robot base frame  $\{F_b\}$  and  $Q_w^i$  in world frame  $\{F_w\} \forall i = 1, \dots, n$ ) is assumed so that

$$\begin{bmatrix} Q_b^i \\ 1 \end{bmatrix} = {}_w^bT \cdot \begin{bmatrix} Q_w^i \\ 1 \end{bmatrix}, \forall i \in I \tag{16}$$

where  ${}_w^bT = \begin{bmatrix} {}_w^bR & {}_w^bt \\ 0 & 1 \end{bmatrix}$  is the transformation matrix between the robot base frame and the world frame such that  ${}_w^bR = [{}_w^br_1 \ {}_w^br_2 \ {}_w^br_3]$ . Expanding Eq. (16) and stacking it for each  $n$  calibration point  $Q^i (I = 1, 2, \dots, n)$  gives a system of linear equations

$$A_T \cdot t = \tag{17}$$

where  $A_T$  is a  $3n \times 12$  coefficient matrix and  $t$  is a  $3n \times 1$  column vector;  $t = \begin{bmatrix} {}_w^br_1 \\ {}_w^br_2 \\ {}_w^br_3 \\ {}_w^bt \end{bmatrix}$ . The transformation matrix  ${}_w^bT$

can be estimated by solving Eq. (17). A summary of these calibrations are tabulated in Table 2. The column ‘‘Method’’ lists the methods to obtain the matrix in column ‘‘Outcome’’ using the equations listed in the ‘‘Equation’’ column.

## 4 Calibration

The autonomous kiwifruit harvesting system has a stereo vision system of two cameras as the sensor for localising kiwifruit, and a robotic arm for picking the kiwifruit as shown in Fig. 4.

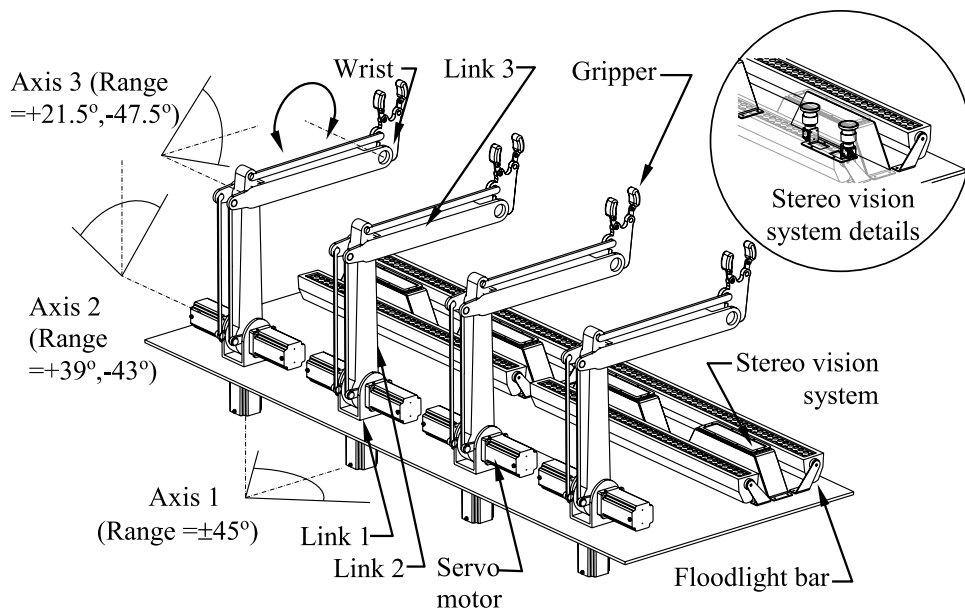
The manipulator consists of three major links and a wrist. These links are actuated by three servo motors installed at the base. At the HOME position of the robot, link 2 aligns with link 1 vertically. Link 3 is positioned horizontally, and the gripper is oriented vertically upward as shown in the figure. Figure 5 shows a schematic configuration of the kiwifruit harvesting robot.  $\theta_1, \theta_2$  and  $\theta_3$  are the joint parameters driving the link 1, 2 and 3 with link length are  $l_1, l_2$  and  $l_3$  respectively. Parameter  $a_j, b_j$  and  $c_j (j = 2, 3)$  are link lengths of the two four-bar linkages.

The orientation of the gripper  $\theta_4 = \pi - (\varphi - \theta_3)$ . The coordinates of pick point  $Q$  are expressed in the robot base frame  $F_b$  as

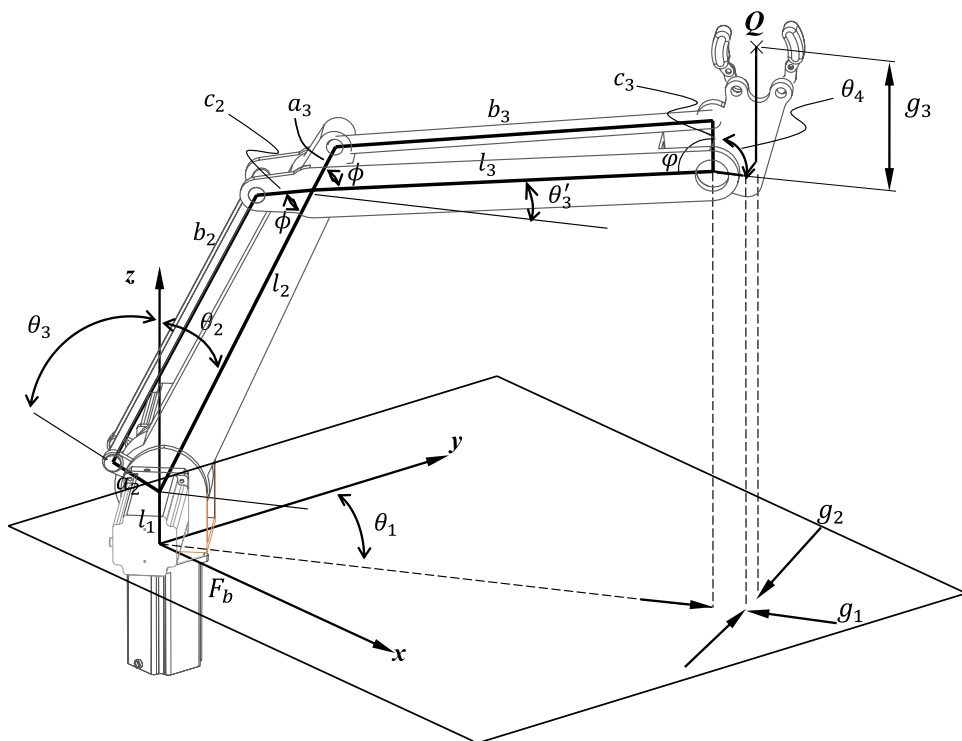
**Table 2** A summary of calibrations

Calibration	Method	Outcome	Equation
Camera calibration	A set of $n$ calibration points	$H$	(4)
	A set of $m$ views	$D$	(7)
	Cholesky decomposition	$K$	(8)
	Computation	${}_r^wT$	(9)
Stereo vision system calibration	A set of $n$ calibration points	$F$	(11)
	Computation	$E$	(12a)
	Single value decomposition	${}_l^rT$	(12b)
Kinematic calibration	A set of $n$ calibration points	$\Delta \bar{a}$	(15)
Integration calibration	A set of $n$ calibration points	${}_w^bT$	(17)

**Fig. 4** The autonomous kiwi-fruit harvesting system



**Fig. 5** A kiwifruit robot arm manipulator with two four bar linkages



$$Q_b = f(\mathbf{g}, \boldsymbol{\theta}) = \begin{bmatrix} (l_2 \sin \theta_2 + l_3 \cos \theta'_3 + g_2) \sin \theta_1 + g_1 \cos \theta_1 \\ (l_2 \sin \theta_2 + l_3 \cos \theta'_3 + g_2) \cos \theta_1 + g_1 \sin \theta_1 \\ l_1 + l_2 \cos \theta_2 + l_3 \sin \theta'_3 + g_3 \end{bmatrix} \in R^3 \quad (18)$$

where  $f(\mathbf{g}, \boldsymbol{\theta})$  is the kinematic model of the robot with  $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} \in R^3$  and  $\mathbf{g} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} \in R^3$ .

$$\theta'_3 = \frac{\pi}{2} - \theta_2 - \phi$$

$$\phi = \sin^{-1} \left( \frac{a_2 \sin(\theta_2 - \theta_3)}{e_2} \right) + \cos^{-1} \left( \frac{e_2^2 + c_2^2 - b_2^2}{2e_2 c_2} \right)$$

$$e_2^2 = a_2^2 + l_2^2 - 2a_2 l_2 \cos(\theta_2 - \theta_3)$$

$$\varphi = \sin^{-1}\left(\frac{a_3 \sin \phi}{e_3}\right) + \cos^{-1}\left(\frac{e_3^2 + c_3^2 - b_3^2}{2e_3 c_3}\right)$$

$$e_3^2 = a_3^2 + l_3^2 - 2a_3 l_3 \cos \phi$$

The angle  $\theta'_3$ ,  $\phi$  and  $\varphi$  are shown in Fig. 5.

The stereo vision system has two Basler ac1920-40uc USB 3.0 cameras with Kowa lens LM6HC F1.8 f6mm 1" format. The devices are installed in a way that two pairs of stereo cameras share two pairs of floodlight bars. Each consists of a 200 kW floodlight bar with 40 LEDs which illuminates the canopy with white light. A 3-D space measurement system is required to perform measurements that will quantify the magnitude of inaccuracies contributed by the stereo vision system and robot of the harvesting system.

Appropriate calibrations can be applied once these inaccuracies have been quantified. Figure 6a, b depict the setup of the measurement system. It consists of a frame holding a constraint plate which is above the robot and the stereo vision system. The constraint plate is connected to a sub-frame through a ball head so that its position and orientation are adjustable. Its height  $h$  is also adjustable using the adjusting unit. The constraint plate is made from an aluminium-acrylic composite sandwich material which has a flatness of

0.3 mm while suspended. A permanent marker and a dial gauge with resolution of 10  $\mu\text{m}$  are installed at the end of the manipulator as shown in Fig. 6c. A small spring is placed under the marker so that it can be pressed against the constraint plate. The amount of compression is measured by the dial gauge. As the manipulator is moved to a measurement position, it leaves a mark indicating the marker position. The spring action of the marker is implemented so that the marker tip position can either be varied (higher or lower) across the constraint plate. Hence, the position of the marker can be measured with reference to the frame.

Various calibration processes can be performed by attaching different calibration tools such as a checkerboard or a board with a grid of dots to the constraint plate. For instance, a checkerboard is used for camera calibration as depicted in Fig. 7a, b shows a board with a grid of dots attached to the constraint plate for stereo vision system calibration, kinematic calibration and integration calibration. The world frame (in red) is set at the centre of the calibration tool as shown.

As the robot moves back and forth, dots are drawn on the board fixed with the constraint plate. The board can then be removed from the frame and placed on a bench to measure the dot positions. Measurements are made by measuring the deviation in the  $x$  and  $y$  directions of the dots

Fig. 6 The measurement system

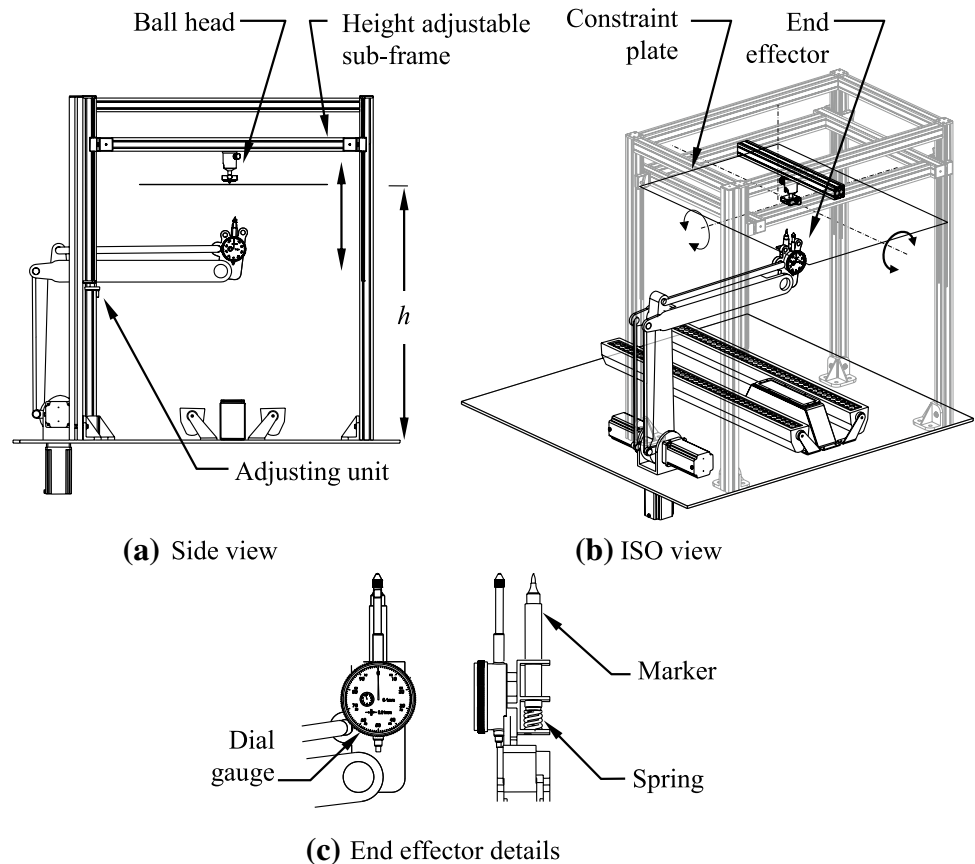
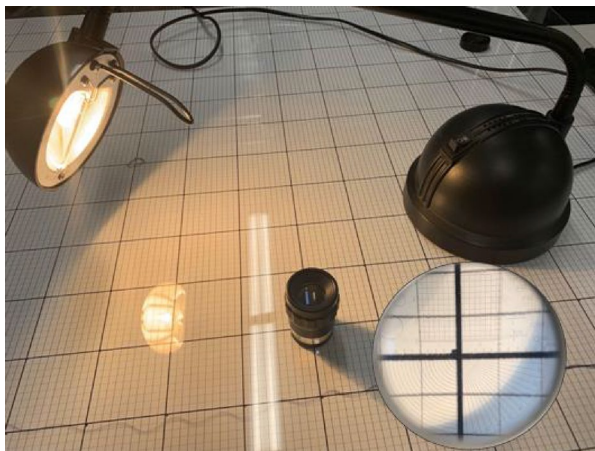
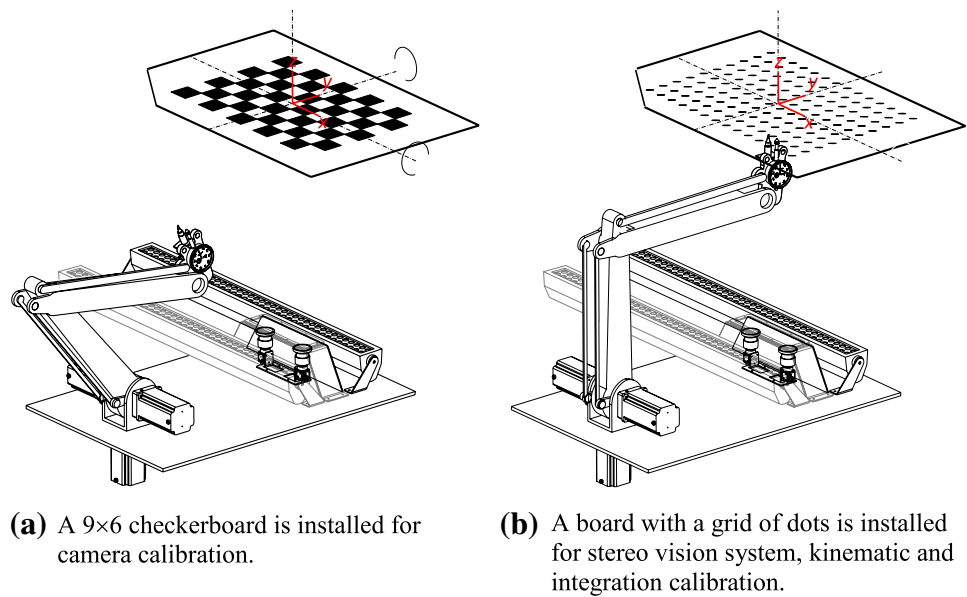
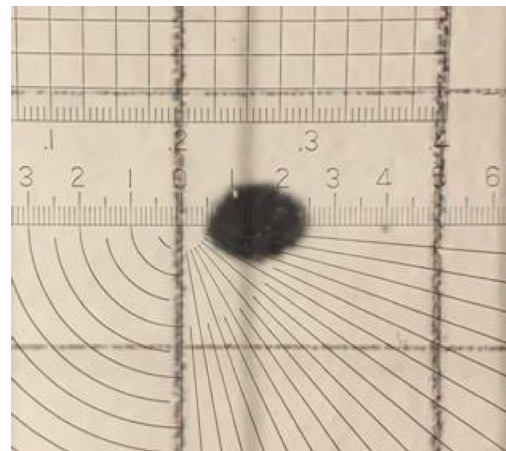


Fig. 7 The calibration tools



(a) Optical magnifier



(b) Example

Fig. 8 Dot position measurement

from their target positions. A 10× etched optical magnifier with an etched ruler with 100 μm increments as shown in Fig. 8a is used during the measurement process to get sub-millimetre resolution accuracy. An example measurement is shown in Fig. 8b (the optical magnifier is not quite aligned with the grid in this image as it is difficult to both align the magnifier and take an image simultaneously). The centre of the dot shown is estimated by measuring the position of the outside edges of the dot and using them to calculate a centre-point. In this case the centre of the dot would be measured as 1.5 mm away from the reference line. The system performs dot position measurements for verifying the kinematic calibration and integration calibration.

#### 4.1 Camera calibration

The cameras in the stereo vision system are calibrated individually to obtain their intrinsic and extrinsic parameters. A checkerboard (A1—80 mm squares—9×6 vertices, 10×7 squares) is installed on the constraint plate. The cameras are calibrated individually using 50 images taken at multiple orientations, covering the entire field of view of the cameras. The matrix  $H$  and  $D$  are estimated using Eqs. (4) and (7). Moreover, the intrinsic and extrinsic parameters are computed by Eqs. (8) and (9).

### 4.2 Stereo vision system calibration

The checkerboard is replaced by a board with a grid of dots (13×17 dots is spaced 50 mm apart in the  $x$  range of ±400 mm and  $y$  range of ±300 mm) for stereo vision system calibration. Images of the dots are captured by both cameras. The dots are detected using a simple colour filter that generates a black and white mask for both images based on pixels that meet the criteria of the filter. From these masks, the pixel locations of the centroid of the distinguished dots are calculated and are stored as a list of key-points. These key-points are then processed through the undistortion matrix of the stereo vision system and then matched between the left and right images. The fundamental matrix  $F$  and essential matrix  $E$  are then estimated using Eqs. (11) and (12). Consequently, the transformation matrix  ${}_l^rT$  of camera  $L$  with respect to the frame  $\{F_r\}$  of camera  $R$  is obtained from the essential matrix  $E$  by single value decomposition.

### 4.3 Kinematic calibration

The board with a grid of dots is also used for kinematic calibration. The marker is driven by the manipulator so that it moves to each dot on the constraint plane sequentially. The marker leaves a dot at each actual position. The joint parameters  $\theta$  are recorded at each dot. These dots are the actual positions of the marker. The joint parameters are substituted into the kinematic model (Eq. (18)) to obtain the calculated positions. The differences in each pair of calculated position and actual (marked) position of the marker define the position error  $\Delta Q_b$ . The geometric variations are estimated iteratively using Eq. (15).

### 4.4 Integration calibration

An array of dots  $Q_w^i(x_w^i, y_w^i, z_w^i)(\forall i = 1, \dots, n)$  expressed in the world frame  $\{F_w\}$  is used for calibrating the robot base and the stereo vision system integration. The end-effector with the marker is driven to each of these points and their corresponding joint parameters are recorded. The coordinates of this array of dots  $Q_b^i(x_b^i, y_b^i, z_b^i)(\forall i = 1, \dots, n)$  expressed in the robot base frame  $\{F_b\}$  are obtained by substituting the joint parameters into the calibrated kinematic model as modelled by Eq. (18). The transformation matrix  ${}_w^bT$  is computed by solving Eq. (17).

### 4.5 Solving a system of linear equations

All the calibrations result in solving a system of equations such as Eqs. (4), (7), (11), (15) and (17) in the forms of

$$A \cdot x = b \tag{19}$$

When  $b = 0$  such as Eqs. (4), (7) and (11), the solving of this system of homogeneous equation can be formulated as estimating  $\hat{x}$  by

$$\min_x (\hat{x}^T A^T \cdot A \hat{x}) \text{ such that } \hat{x}^T \cdot \hat{x} = 1 \tag{20}$$

A Loss function  $L(\hat{x}, \lambda)$  is defined such that

$$L(\hat{x}, \lambda) = \hat{x}^T A^T \cdot A \hat{x} - \lambda (\hat{x}^T \cdot \hat{x} - 1) \tag{21}$$

Taking derivative of  $L(\hat{x}, \lambda)$  with respect to  $\hat{x}$  yields an eigen value problem

$$A^T \cdot A \hat{x} = \lambda \hat{x} \tag{22}$$

$\hat{x}$  is an eigen vector with smallest eigen value  $\lambda$  of matrix  $A^T \cdot A$  minimizing the Loss function  $L(\hat{x}, \lambda)$ . Once variable  $\hat{x}$  is estimated, the corresponding matrix can be obtained by re-arranging the components of variable  $\hat{x}$ .

If  $b \neq 0$ , the equations are linear such as Eqs. (15) and (17), the least squares solution  $\hat{x}$  is estimated

$$\hat{x} = (A^T A)^{-1} A^T b \tag{23}$$

However, Eq. (15)  $A_k \cdot \Delta \bar{a} = \Delta \bar{Q}_b$  in kinematic calibration is a non-linear system and  $\Delta \bar{a}$  is computed by iteration until convergence. At the  $j$ -th iteration,  $\Delta \bar{a}^j$ ,  $\Delta \bar{Q}_b^j$  and  $A_k$  are denoted as  $\Delta \bar{a}^j$ ,  $\Delta \bar{Q}_b^j$  and  $A_k^j$  ( $A_k^j$  contains  $\bar{a}^j$  and  $\theta$ , the joint parameters for  $n$  calibration points) respectively,

$$A_k^j = A_k^j(\bar{a}^j, \theta) \tag{24a}$$

$$\Delta \bar{a}^j = \left( (A_k^j)^T \cdot A_k^j \right)^{-1} \cdot (A_k^j)^T \cdot \Delta \bar{Q}_b^j \tag{24b}$$

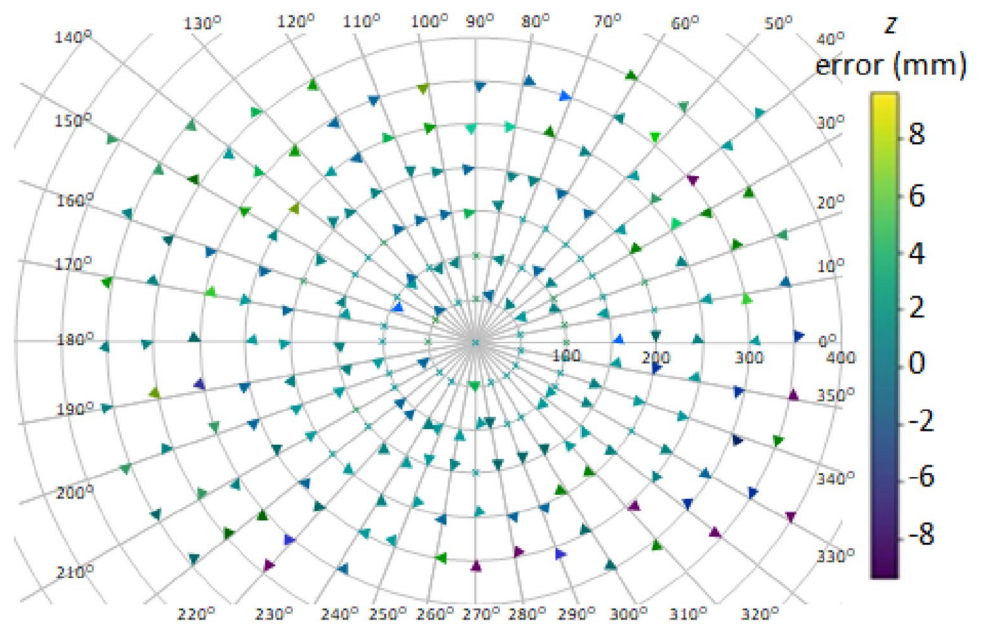
$$\bar{a}^{j+1} = \bar{a}^j + \Delta \bar{a}^j \tag{24c}$$

$$\Delta \bar{Q}_b^{j+1} = A_k^j \cdot \Delta \bar{a}^j \tag{24d}$$

## 5 Verification of calibration processes

A constraint plate with a circular array of dots is used for accuracy verification. These dots are arranged in polar coordinates with the origin at the home position of the robot. The verifications are conducted at three height level  $h$  of 700 mm, 800 mm and 900 mm. The measured dots are plotted with the circular array of dots for comparison. The plot is shown as arrows pointing from the measured dot positions to the dot positions in the circular array. The length of the

**Fig. 9** Stereo vision system accuracy of the kiwifruit harvesting system after calibration



**Table 3** Summary of stereo vision system accuracy before and after calibration

	Average error	Maximum error	Error range
$r$ (mm)	-1.02	-3.15	5.65
$\theta$ (°)	1.52	2.54	3.41
$z$ (mm)	-1.75	-8.2	14.8

arrow is proportional to the error magnitude in radial and circumferential direction. The colour of the arrow is based on the measured  $z$  error, which is the depth of that data point, with purple being below the average depth and yellow being above.

In the stereo vision system verification, the sensed dot (measured dots) locations expressed in the world frame  $\{F_w\}$  are calculated. The discrepancies between the sensed dots and the dots in the circular array are plotted in Fig. 9. The summary of the plot is tabulated in Table 3.

In the kinematic accuracy verification, the end-effector is driven to each point of the circular array. The joint parameters are recorded. The calculated end-effector location for each set of joint parameters is obtained from Eq. (18). The positional errors between these calculated locations and their corresponding dot locations in the circular array are plotted in Fig. 10. Figure 10a, b show the positional errors before and after the kinematic calibration. The summary of the plot is tabulated in Table 4.

The accuracy of integration calibration is verified by considering the hand-eye coordination accuracy of the autonomous harvesting system. A circular array of dots is

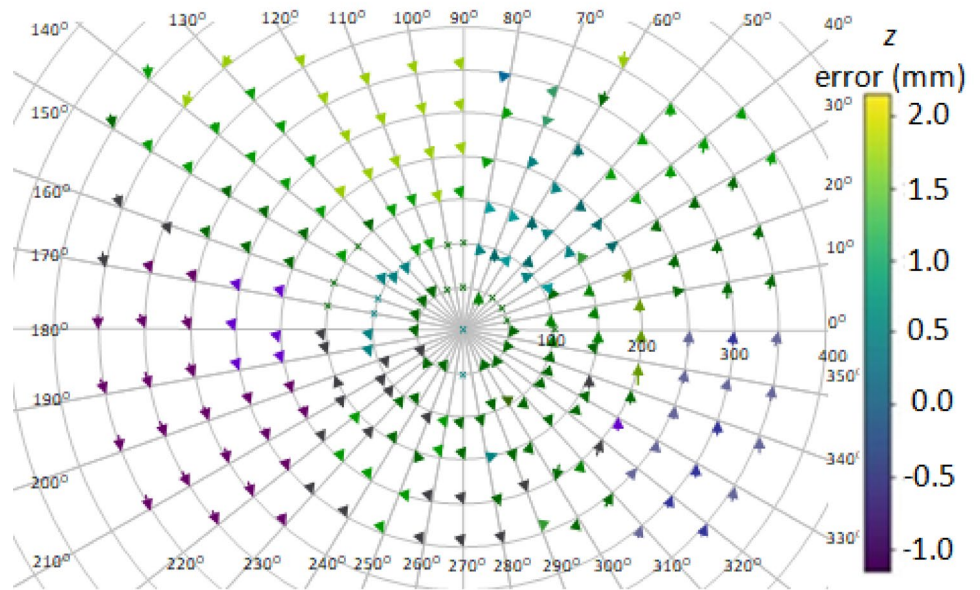
sensed by the stereo vision system. The coordinates of the dots  $Q_r^i$  are computed and expressed in the stereo vision system coordinate frame  $\{F_r\}$ . These coordinates are then transformed into the world frame  $\{F_w\}$  to yield coordinates  $Q_w^i$ . By using the transformation matrix  ${}^w_bT$ , these dot coordinates are transformed to  $Q_b^i$  in the robot frame  $\{F_b\}$ . The joint parameters  $\theta^i$  corresponding to each dot coordinate  $Q_b^i$  are obtained by inverse kinematics. Eventually, the end-effector is driven by these joint parameters  $\theta^i (\forall i = 1, \dots, n)$  to mark the dots on the constraint plate. The differences between the original circular array of dots and the marker positions are plotted in Fig. 11. The results are summarized in Table 5.

## 6 Lab tests

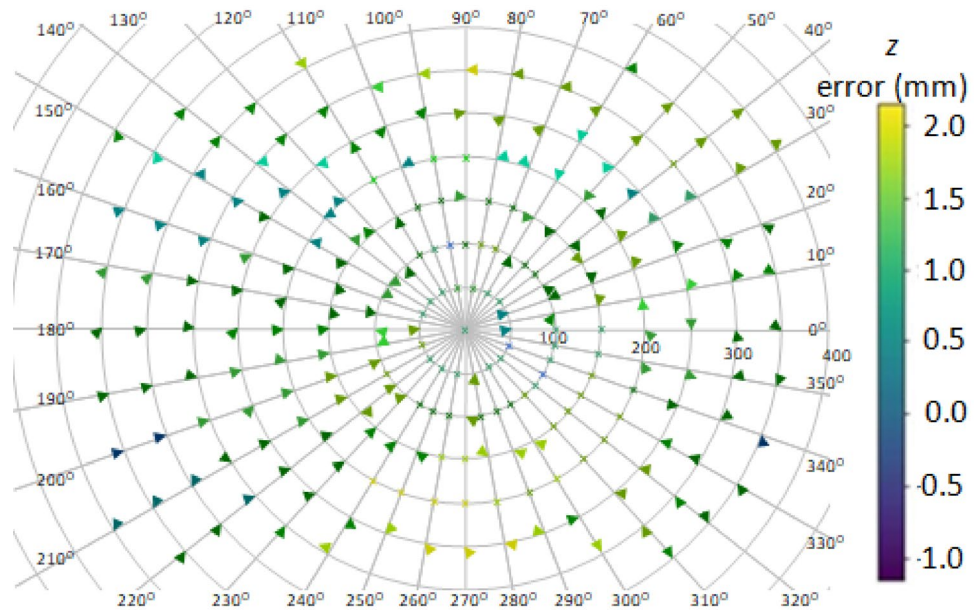
Several tests are set up to check the performance of the autonomous kiwifruit harvesting system after calibrations using the fake kiwifruit in the laboratory as shown in Fig. 12a. As the kiwifruit tends to grow in cluster, the fake kiwifruit is arranged as individual fruit, three fruit clusters, two fruit cluster and seven fruit cluster as shown in Fig. 12b–e.

The autonomous kiwifruit harvesting system is placed under the fruit which are hung on the lab ceiling. The fruit is localized by the stereo vision system and the corresponding joint parameters of the manipulator are then computed to drive the gripper in picking the fake kiwifruit. The outcomes are categorized into three categories of picked, dropped and unpicked as tabulated in Table 6.

**Fig. 10** Kinematic accuracies of the kiwifruit harvesting system before and after calibration



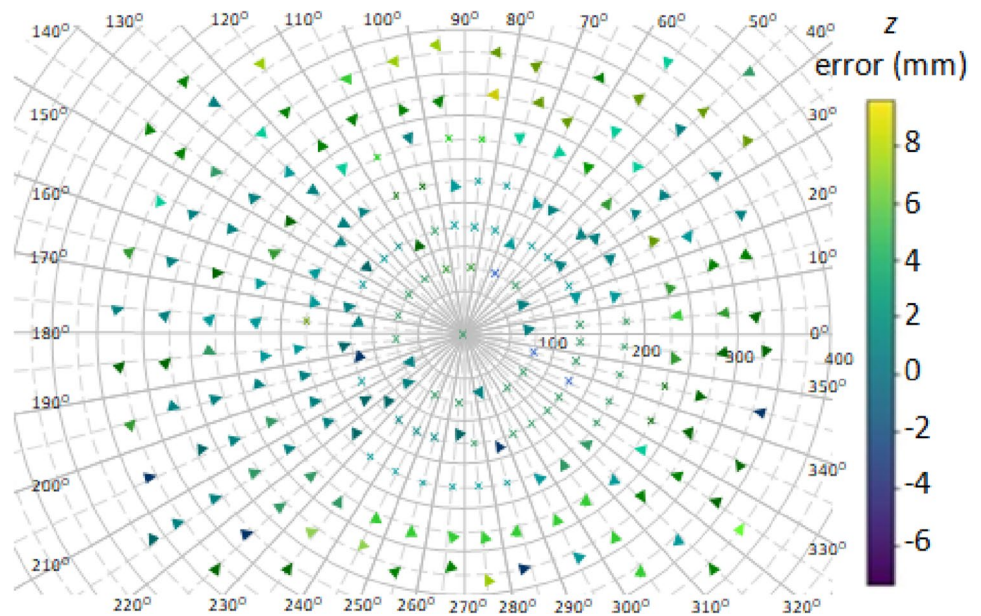
**(a)** Un-calibrated kinematic accuracy.



**(b)** Calibrated kinematic accuracy.

**Table 4** Summary of kinematic accuracy before and after calibration

	Average error		Maximum error		Error range	
	Un-calibrated	Calibrated	Un-calibrated	Calibrated	Un-calibrated	Calibrated
$r$ (mm)	4.92	1.71	12.47	2.56	20.5	2.75
$\theta$ (°)	3.54	1.35	- 5.5	3.23	7.78	2.2
$z$ (mm)	- 0.76	0.91	2.15	2.06	3.05	2.25

**Fig. 11** The accuracy of hand–eye coordination**Table 5** Summary of the hand–eye coordination accuracy

	Average error	Maximum error	Error range
$r$ (mm)	0.7	3.15	5.65
$\theta$ °	1.52	2.54	3.41
$z$ (mm)	1.95	7.12	11.8

## 7 Orchard tests

Orchard tests are conducted in the several commercial Hayward orchards to determine the performance of the calibrated harvesting system as shown in Fig. 13.

Four bays of kiwifruit are harvested to determine the performance. Each bay is divided into several zones. At each zone, the harvesting system is instructed to pick the fruit that are visible by the stereo vision system. At the end of harvesting in a zone, fruit harvested and fruit dropped are manually picked and counted. Fruit left in the canopy is also manually harvested and counted. The harvesting processes are video recorded so that a more detailed analysis can be performed later. The results are summarized in Table 7 which are compared to the performance before calibration.

## 8 Discussion

The calibrations can be categorized into two parts. The stereo vision calibration and kinematic calibration. Stereo vision techniques use two cameras to view the same object. The orientations of both cameras installed in the stereo vision system are the same and the distance apart is

known. However, these are hard to guarantee the accuracy due to manufacturing issues, especially the fixture holding the cameras. Such inaccuracies affect the extrinsic parameters of the cameras. The stereo vision system calibration is mainly to evaluate these parameters using a set of calibration points and to eliminate the unwanted noise. Since the cameras are operated under the optical principles, a linear model and a set of linear equations are employed to model the system except the lens distortion. The lens distortion is non-linear and this complicates the calibration process. The effects of lens distortion can be serious if it is not properly undistorted.

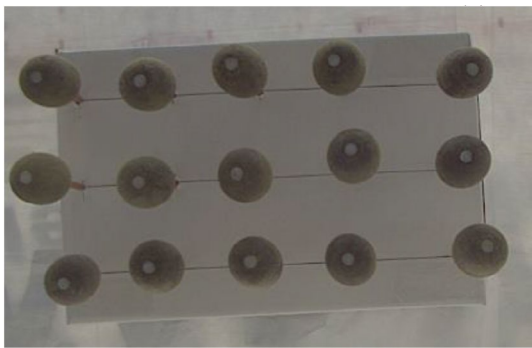
Different from stereo vision system calibration, kinematic calibration is performed to minimize the difference between the desired and actual gripper position. A kinematic model of a manipulator, which depends on its mechanical structure, is required for calibration. The kinematic model is usually a non-linear function due to the revolution joints. As a result, the kinematic calibration results in a system of non-linear equations and iterative processes are required to solve the equations.

In the kinematic calibration accuracy plot, those "crosses" around the plot centre at (0, 0) indicate that the  $r$  and  $\theta$  coordinates of the commanded and actual positions are very close. In addition, the  $z$  coordinate differences are also close to zero. This shows that the assumption of approximating the gripper orientation to be vertical to simplify the inverse kinematic computation is feasible when the gripper is around its home position.

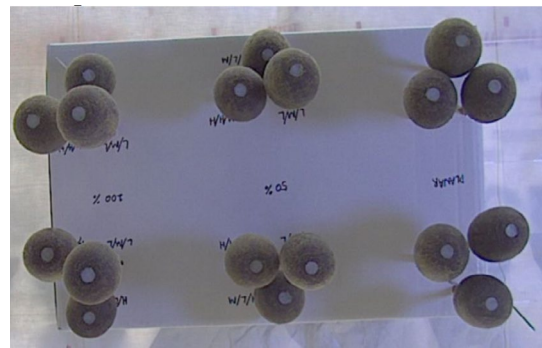
In fact, the plots in Figs. 9, 10, 11 show that the accuracy of the autonomous harvesting system is very high. This also reflects in the laboratory test. The successful picking rate for individual fake kiwifruit is 100%.



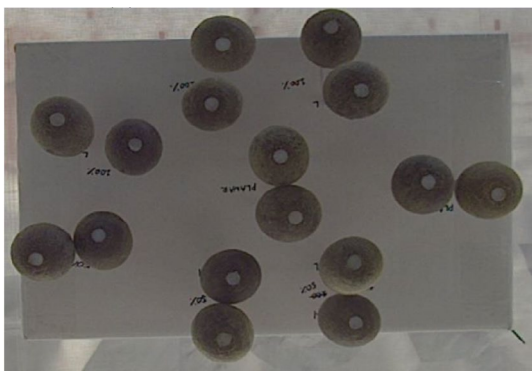
(a) Lab test set up.



(b) Individual fake kiwifruit.



(c) 3 fruit cluster.



(d) 2 fruit cluster.



(e) 7 fruit cluster.

Fig. 12 Laboratory tests of autonomous kiwifruit harvesting system

**Table 6** Results of laboratory tests of autonomous kiwifruit harvesting system

	No. of picks	Picked	Dropped	Unpicked
Individual fruit	Over 300	100%	0%	0%
2 fruit cluster	Over 320	87.5%	6.25%	6.25%
3 fruit cluster	Over 360	57%	19%	23.8%
7 fruit cluster	Over 420	67%	28.57%	4.76%

However, this rate drops when the fruit cluster size increases. The average picking rate drops from 100% for individual fruit to 67% for seven fruit clusters (with a minimum at 57% for three fruit clusters). The actual successful harvesting rate in the orchard test is even lower than that of the laboratory test results. The calibration improves the positioning of the gripper so that the number of entry failure, grip failure, target and non-target knockoffs, and fruit throws and drops are reduced. As a result, the fruit dropped percentage decreases from 24.6 to 10.7% after calibration. However, the successful harvesting rate just rises from 50.8 to 55%. This indicates that those fruits remain in the canopy instead of harvested as the percentage of fruit left

in the canopy increases from 24.6 to 34.3%. This is due to the “offline” stereo vision system for sensing the canopy information, which refers to capturing the canopy once at the beginning of a harvesting cycle.

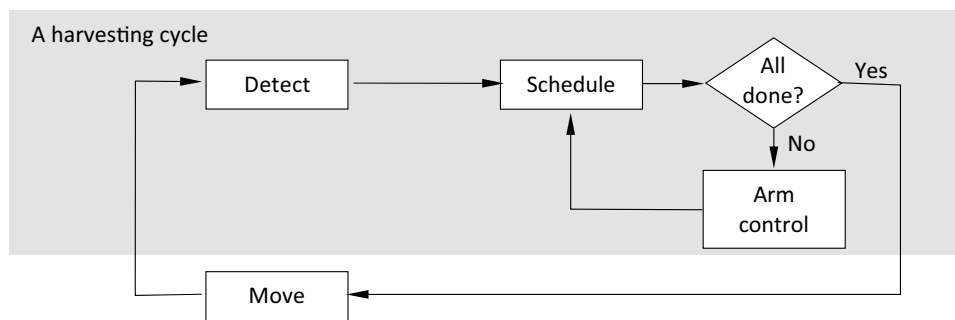
The operation of the autonomous harvesting system at a specific zone under the canopy follows a harvesting cycle as shown in Fig. 14a. The stereo vision system on each module images the canopy. Then the fruits are detected and their positions in the world frame are located. The located fruit positions from each module are collated by a scheduler which assigns the target fruit for each robotic arm dynamically based on the principle of minimum distance with a set of rules for collision prevention. When the robot controller receives a target fruit location from the scheduler, the joint parameters are calculated and the manipulator drives the gripper to perform the harvesting. Then the next target fruit

**Table 7** Orchard test result comparison to before and after calibration

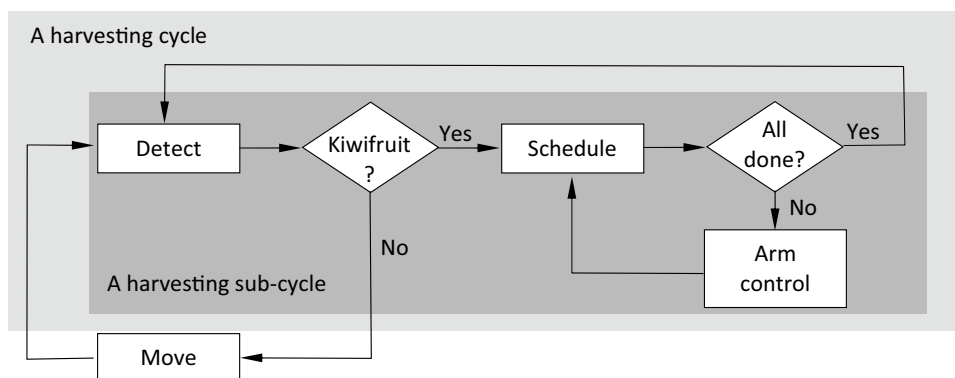
Calibration	No. of picks	Harvested	Dropped	Left
Before	1456	50.8%	24.6%	24.6%
After	5327	55%	10.7%	34.3%

**Fig. 13** Orchard test for autonomous kiwifruit harvesting system

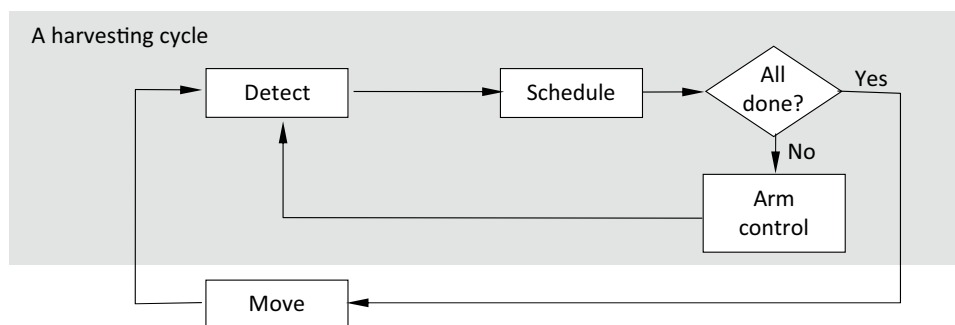
**Fig. 14** The harvesting cycle of the autonomous harvesting system at a specific location



**(a)** The harvesting cycles using the “offline” stereo vision systems.



**(b)** The harvesting cycles using the “offline” stereo vision systems with “sub-cycles”.



**(c)** The harvesting cycles using the “real-time” stereo vision systems.

for the robot will be determined by the scheduler until all target fruit in the scheduler are assigned. Then the autonomous harvesting system is moved to another zone.

Reviewing the harvesting videos shows that most of the kiwifruit grow in clusters. These fruits are not at the same level. Some are at a level higher than the others. Before the calibration, once a fruit is harvested, the gripper usually knocks off the neighbouring fruit because of the position inaccuracy. After calibration, the position error is reduced. The fruit is correctly placed into the gripper and the neighbouring fruit are not knocked off. However, these fruits in the cluster are obstructed by the stereo vision system before

harvesting. They are not visible to the stereo vision system and are not scheduled for harvesting. Hence, they remain in the canopy even though they are not knocked off. As an alternative, the harvesting cycle is re-run as a “sub-cycle” again to pick those fruit left in the canopy at that zone. This sub-cycle repeats until all the fruit are clear from the canopy (Williams et al. 2019). The flow charts of these processes are outlined in Fig. 14b. Hence, the harvesting time is increased depending upon how the kiwifruit grows clusterly.

In order to increase the economic viability of the harvesting system, a real-time stereo vision system (Jin and Tang 2009; Matsumoto and Zelinsky 2000) is proposed

for the next development phase as depicted in Fig. 14c. With the real-time stereo vision system, the instantaneous canopy information is obtained, and the scheduler should update the harvesting order accordingly if an obstructed kiwifruit becomes visible. Furthermore, the target fruit allocation by the scheduler should also be based on evenly workload distribution to minimize the cycle harvesting time. This can be achieved by introducing competition among the robots while they are cooperating to complete the harvesting task. With the real-time stereo vision system, visual servoing control (Chaumette and Hutchinson 2006, 2007) can also be implemented into the manipulator controller to achieve a more accurate fruit harvesting.

## 9 Conclusion

Video analysis is performed to investigate the reasons for high fruit drop rate during harvesting using an autonomous kiwifruit harvesting system. It is revealed that the position error between the gripper and the fruit is the main cause of fruit drop during harvesting. In order to reduce these position errors, an integrated approach for stereo vision system calibration and kinematics calibration is proposed for the autonomous kiwifruit harvesting system.

A measurement system is set up to acquire the data for both calibrations and verifications. Laboratory tests and orchard tests are arranged to investigate the performance of the harvesting system in real application environments. Based on the measurements after these calibrations, it is found that the position error of the harvesting system is greatly reduced. This is actually reflected in the high successful picking rate (100%) in the laboratory test of picking individual fake kiwifruit. The fruit dropped rate is also greatly improved in the real orchard tests because of the position error reduction. However, the successful harvesting rate is not impressively improved as most of the fruit remain unharvested in the canopy even though they are not knocked off due to position error. This is mainly due to the fact that canopy information such as fruit distribution is obtained by the stereo vision system before harvesting starts. These fruits are not harvested because they are obstructed to the stereo vision system as they are hidden by their neighbouring fruit.

As a result, a real-time stereo vision system is required during harvesting to sense the dynamic environment of canopy information to increase the successful harvesting rate and economic viability. Nevertheless, this will complicate the harvesting scheduling and workload distribution among the robots. As the autonomous kiwifruit harvesting system consists of multiple robots, the harvesting task is cooperative and competitive. The robots cooperate to harvest the fruit

and they compete with each other so that the harvesting is completed in a minimum of time.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s41315-022-00263-x>.

**Author contributions** All authors contributed to the study conception and design. Material preparation for Laboratory and orchard test, data collection and analysis were performed by CKA, SHL, YCK, MR and CT. The first draft of the manuscript was written by CKA and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions. The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

**Data availability** The custom code and data used during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Ethics approval** This study does not require ethics approval as no human or animal subjects are involved.

**Consent to participate** This study does not require consent to participate as no human research participant is included in the study.

**Consent to publish** Authors consent to publish the here presented work.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Brown, D.C.: Close-range camera calibration. *Photogramm. Eng.* **37**(8), 855–866 (1971)
- Chaumette, F., Hutchinson, S.: Visual servo control. I. Basic approaches. *IEEE Robot. Autom. Mag.* **13**(4), 82–90 (2006). <https://doi.org/10.1109/MRA.2006.250573>
- Chaumette, F., Hutchinson, S.: Visual servo control. II. Advanced approaches. *IEEE Robot. Autom. Mag.* **14**(1), 109–118 (2007). <https://doi.org/10.1109/MRA.2007.339609>
- Cui, J., Huo, J., Yang, M.: Novel method of calibration with restrictive constraints for stereo-vision system. *J. Mod. Opt.* **63**(9), 835–846 (2016). <https://doi.org/10.1080/09500340.2015.1106602>

- Feng, X., Fang, B.: Algorithm for epipolar geometry and correcting monocular stereo vision based on a plane mirror. *Optik* (2021). <https://doi.org/10.1016/j.ijleo.2020.165890>
- Jin, J., Tang, L.: Corn plant sensing using real-time stereo vision. *J. Field Robot.* **26**(6–7), 591–608 (2009). <https://doi.org/10.1002/rob.20293>
- Lightcap, C., Hamner, S., Schmits, T., Banks, S.: Improved positioning accuracy of the PA10-6CE robot with geometric and flexibility calibration. *IEEE Trans. Rob.* **24**(2), 452–456 (2008). <https://doi.org/10.1109/TRO.2007.914003>
- Liu, J.Z., Li, P.P., Mao, H.P.: Mechanical and kinematic modeling of assistant vacuum sucking and pulling operation of tomato fruits in robotic harvesting. *Trans. ASABE* **58**(3), 539–550 (2015)
- Matsumoto, Y. and Zelinsky, A.: An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition* (Cat. No. PR00580), 28–30 March 2000 (2000) <https://doi.org/10.1109/AFGR.2000.840680>
- Memon, Q., Khan, S.: Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks. *Int. J. Syst. Sci.* **32**(9), 1155–1159 (2010). <https://doi.org/10.1080/00207720010024276>
- Moré, J.J.: The Levenberg–Marquardt algorithm: implementation and theory. In: Watson, G.A. (ed.) *Numerical analysis Lecture Notes in Mathematics*, vol. 630. Springer, Berlin (1987). <https://doi.org/10.1007/BFb0067700>
- Motta, J., de Carvahó, G., McMaster, R.S.: Robot calibration using a 3D vision-based measurement system with a single camera. *Robot. Comput.-Integr. Manuf.* **17**(6), 487–497 (2001). [https://doi.org/10.1016/S0736-5845\(01\)00024-2](https://doi.org/10.1016/S0736-5845(01)00024-2)
- NZ MPI. MPI helps attract people to pick green gold on kiwifruit orchards. <https://www.mpi.govt.nz/news/media-releases/mpi-helps-attract-people-to-pick-green-gold-on-kiwifruit-orchards/>. (2021)
- NZ MSD. Bay of Plenty Labour Shortage Declaration. Retrieved September 24, 2018, from <https://www.msd.govt.nz/about-msd-and-our-work/newsroom/media-releases/2018/bay-of-plenty-labour-shortage-declaration.html>. (2018)
- Mu, L., Cui, G., Liu, Y., Cui, Y., Fu, L.: Design and simulation of an integrated end-effector for picking kiwifruit by robot. *Inf Process Agric* **7**, 58–71 (2020)
- Nubiola, A., Bonev, I.A.: Absolute calibration of an ABB IRB 1600 robot using a laser tracker. *Robot. Comput.-Integr. Manuf.* **29**(1), 236–245 (2013). <https://doi.org/10.1016/j.rcim.2012.06.004>
- NZKGI. New Zealand Kiwifruit Labour Shortage. <https://nzkgi.org.nz/wp-content/uploads/2018/07/NZ-Kiwifruit-Labour-Shortage-July-2018.pdf>. (2018)
- NZKGI. NZKGI Labour Report for the 2020 Season. <https://www.nzkgi.org.nz/wp-content/uploads/2021/10/NZKGI-Labour-Report-for-the-2020-season.pdf#article>. (2020)
- Park, K., Park, C. and Shin, Y.: Performance Evaluation of Industrial Dual-Arm Robot. 2008 International Conference on Smart Manufacturing Application. (2008) <https://doi.org/10.1109/ICSMA.2008.4505596>
- Rocha, C.R., Tonetto, C.P., Dias, A.: A comparison between the Denavit–Hartenberg and the screw-based methods used in kinematic modeling of robot manipulators. *Robot. Comput.-Integr. Manuf.* **27**(4), 723–728 (2011). <https://doi.org/10.1016/j.rcim.2010.12.009>
- Silwal, A., Davidson, J.R., Karkee, M., Mo, C., Zhang, Q., Lewis, K.: Design, integration, and field evaluation of a robotic apple harvester. *J. Field Robot.* **34**(6), 1140–1159 (2017)
- Statistics New Zealand. Minimum wage pushes up private sector pay rates. Retrieved September 24, 2018, from <https://www.stats.govt.nz/news/minimum-wage-pushes-up-private-sector-pay-rates>. (2018)
- Tsai, R.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE J. Robot. Autom.* **3**(4), 323–344 (1987)
- Weng, J., Cohen, P., Hemiou, M.: Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(10), 965–980 (1992). <https://doi.org/10.1109/34.159901>
- Williams, H.A., Jones, M.H., Nejati, M., Seabright, M.J., Bell, J., Penhall, N.D., Barnett, J.J., Duke, M.D., Scarfe, A.J., Ho, S.A., Lim, J.Y., MacDonald, B.A.: Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms. *Biosys. Eng.* **181**, 140–156 (2019). <https://doi.org/10.1016/j.biosysteng.2019.03.007>
- Williams, H.A., Ting, C., Nejati, M., Jones, M.H., Penhall, N., Lim, J.Y., Seabright, M., Bell, J., Ahn, H.S., Scarfe, A., Duke, M., MacDonald, B.A.: Improvements to and large-scale evaluation of a robotic kiwifruit harvester. *J. Field Robot.* **37**(2), 187–201 (2020). <https://doi.org/10.1002/rob.21890>
- Xiong, Y., Ge, Y., Grimstad, L., From, P.J.: An autonomous strawberry-harvesting robot: design, development, integration, and field evaluation. *J. Field Robot.* **37**, 202–224 (2020). <https://doi.org/10.1002/rob.21889>
- Xu, W. and Mills, J. K.: A new approach to the position and orientation calibration of robots. In *Proceedings of the 1999 IEEE International Symposium on Assembly and Task Planning (ISATP'99)* (Cat. No.99TH8470) (1999) <https://doi.org/10.1109/ISATP.1999.782970>
- Zespri. *Zespri Annual Report 2020/21*. [https://www.zespri.com/content/dam/zespri/nz/annual-reports/Annual\\_Report\\_2020\\_21.pdf](https://www.zespri.com/content/dam/zespri/nz/annual-reports/Annual_Report_2020_21.pdf). (2020)
- Zhang, Z.: Determining the epipolar geometry and its uncertainty: a review. *Int. J. Comput. Vision* **27**, 161–195 (1998). <https://doi.org/10.1023/A:1007941100561>
- Zhang, S.: *Handbook of 3D Machine Vision, Optical Metrology and Imaging*. CRC Press, Boca Raton (2016)
- Zhang, H., Zhang, L., Wang, H., Chen, J.: Surface measurement based on instantaneous random illumination. *Chin. J. Aeronaut.* **22**(3), 316–324 (2009). [https://doi.org/10.1016/S1000-9361\(08\)60105-3](https://doi.org/10.1016/S1000-9361(08)60105-3)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Chi Kit Au** received his PhD from Hong Kong University of Science and Technology in mechanical engineering. He is a senior lecturer in University of Waikato. His research interests are CAD/CAM and robotics.



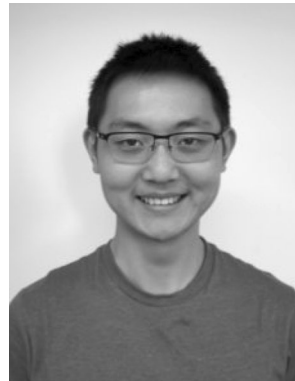
**Shen Hin Lim** received his Ph.D. from University of New South Wales, Sydney in 2010. He is currently a senior lecturer in University of Waikato focussing on smart system development and control in horticulture field, and agri-robotics.



**Michael Redstall** received his master of engineering in electronic engineering from University of Waikato. His research interest is robotics.



**Professor Mike Duke** is the Acting Dean and Dr. John Gallagher Chair in Engineering at the University of Waikato. He is a founding member of Waikato Robotics Automation and Sensing research group and he is active in New Zealand's research landscape.



**Canaan Ting** received his master of engineering in mechanical engineering from University of Waikato. His research interest is robotics.



**Ye Chow Kuang** received the B.Eng. degree (Hons.) in electromechanical engineering and the Ph.D. degree in non-invasive diagnostic techniques from the University of Southampton. He worked in the areas of machine intelligence, uncertainty modelling in engineering design, and automation. He is currently affiliated to the Waikato Robotic, Automation and Sensing Group as well as Artificial Intelligence Institute of the University of Waikato.