Available online at www.sciencedirect.com

ScienceDirect

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

Research Paper

Symmetry-based 3D shape completion for fruit localisation for harvesting robots

Yuanyue Ge [a,b], Ya Xiong [a,*], Pål J. From [a]

[a] Faculty of Science and Technology, Norwegian University of Life Sciences, Ås, Norway
[b] School of Mechanical and Electrical Engineering, Chuzhou University, Chuzhou, China

ARTICLE INFO

Article history:
Received 14 December 2019
Received in revised form
28 June 2020
Accepted 3 July 2020
Published online 21 July 2020

Keywords:
Strawberry harvesting
Machine vision
Localisation
Shape completion

Fruit localisation is a crucial step in developing a robotic fruit-harvesting system. This paper aims to improve the localisation accuracy of fruits in 3D space. In the machine vision system of a harvesting robot, in a single view the visible area of a target is often incomplete and therefore, cannot be directly used to accurately determine the target location. A 3D shape completion method is proposed that can be used on the partially visible images of strawberries obtained from a single view. This method proposed a given number of symmetric plane candidates based on the assumption that the targets are symmetrical, which is normally true for fruits such as such apples, citrus fruits and strawberries. Corresponding rating rules were proposed to select the optimal symmetry to be used for the shape completion. The algorithm was then tested on reconstructed point clouds and implemented on a strawberry harvester equipped with a Red Green Blue-Depth (RGB-D) camera. The evaluation on reconstructed strawberry data showed that the intersection over union (IoU) and centre deviation between the results obtained by this method and ground truth were 0.77 and 6.9 mm, respectively, whilst those of the unprocessed partial data were 0.56 and 14.1 mm. The evaluation results of the strawberry data captured with the RGB-D camera showed that the IoU and centre deviation between the results obtained by this method and ground truth were 0.61 and 5.7 mm, respectively, whilst those of the unprocessed partial data were 0.47 and 8.9 mm.
© 2020 The Author(s). Published by Elsevier Ltd on behalf of IAgrE. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The shortage of human pickers and increasing labour costs have led to a high demand for the automation in fruit harvesting. Many research projects on agricultural robots are engaged in various aspects of agricultural automation, such as apple and sweet pepper harvesting and wine grape harvesting (Bac et al., 2017; Lehnert, English, McCool, Tow, & Perez, 2017; Reis et al., 2012; Silwal, Davidson, Karkee, Zhang, Lewis, 2017). However, challenges remain regarding the development of a robust and commercially available robot, which include, but are not limited to, target detection and localisation (Gongal, Amatya, Karkee, Zhang, & Lewis,

* Corresponding author.
E-mail address: caucoexy@hotmail.com (Y. Xiong).
https://doi.org/10.1016/j.biosystemseng.2020.07.003
1537-5110/© 2020 The Author(s). Published by Elsevier Ltd on behalf of IAgrE. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Nomenclature

Symbols

c	A constant to determine the point position on the line
Dist_ref	Reference distance for the size of the strawberry
Dist_tol	A tolerant distance
Dist2c	Distance between mirrored point and the detected strawberry centre
Dist2c_0	radii of a circle to define point position
Dist2c_1	radii of a circle to define point position
j	The number of rotations along x
j_1	A constant to define value of *Dist_tol*
j_2	A constant to define value of *Dist2c_0*
j_3	A constant to define value of *Dist2c_1*
k	The number of rotations along y
k_1-k_5	Rating adjustment constant
m	The number of translations along depth
Mp	Mask of detected strawberry
Mp'	Projected mask of mirrored point cloud P'
$\boldsymbol{n}_0$	A normal vector that is perpendicular to the initial hypothetical plane
$\boldsymbol{n}_{new}$	A normal vector that is perpendicular to the new hypothetical plane
p	A point in the detected strawberry point cloud
p'	The mirrored point of p
P	Original point cloud of the detected strawberry
P'	Mirrored point cloud of P
p^H	Any point on the initial hypothetical plane
p_0^H	A point on the initial hypothetical plane
p^{HL}	A point on both the line p^L and the hypothetical plane
p^L	A line (any point on the line) passes though p_0^L and perpendicular to the hypothetical plane
p_0^L	A point in the detected strawberry point cloud and on the line p^L
p'_n	De-projected 3D point of p'_{nm}
p'_m	Projected point of p'
p'_{nm}	Nearest point of p'_m in Mp
R	Rotation matrix in 3D
$\boldsymbol{Vec2n}$	Vector from mirrored point p' to nearest point p'_n
Δd	A regular distance interval along depth direction
$\Delta\theta_1$	A regular rotation angle interval along x axis
$\Delta\theta_2$	A regular rotation angle interval along y axis

Abbreviations

CCD	Charge coupled device
IoU	Intersection over union
RGB-D	Red Green Blue-Depth
TOF	Time of flight

2015), as well as end effector design and manipulation algorithm development. Localisation refers to the goal of detection in harvesting vision system and is essential to further fruit manipulation. Therefore, the accurate localisation of fruits along with the effective manipulation algorithms are key factors to determine the performance of robotic harvesting.

In a precision agricultural systems, localisation can be achieved in different ways, for example using a monocular vision system with other predefined conditions or methods (De-An, Jidong, Wei, Ying, & Yu, 2011; Mehta & Burks, 2014; Reis et al., 2012; Xiong, Ge, Liang, & Blackmore, 2017), stereo cameras (Bac, Hemming, & Van Henten, 2014; Ji, Meng, Qian, Xu, & Zhao, 2017; Mehta & Burks. 2016) and other 3D cameras (Vitzrabin & Edan, 2016; Wang, Walsh, & Verma, 2017; Xiong, Peng, Grimstad, From, 2019). Several researchers and practitioners are starting to favour RGB-D cameras because of their simplicity and high localisation accuracy. However, regardless of the camera used, the harvesting system can only obtain partial target information using a single camera from a single view, whilst 3D reconstructions can acquire complete information, but it usually requires a scanning motion from the camera and is therefore computational expensive and not practical in a real-time harvesting system. However, a complete shape of a target fruit is essential for localisation in 3D space, especially when the targets are clustered together, such as with strawberries.

Therefore, here the aim was to improve localisation accuracy by completing the target position for partially visible targets. The complete shape information can be obtained with 3D scanning methods (Le Cozler et al., 2019; Mack et al., 2018; Mack, Lenz, Teurine, Steinhage, 2017). However, as described above, these are not viable choices in harvesting systems. Researchers are working on alternative methods to recover the entire shape of the targets, especially for the purpose of grasping (Bohg et al., 2011; Makhal, Thomas, & Gracia, 2018; Schiebener, Schmidt, Vahrenkamp, & Asfour, 2016). The most common method for completing the shape of a fruit by proposing symmetry was initiated by Thrun and Wegbreit (2005).

The scenario involved in most shape completion methods is the presence of a plane table with objects in a laboratory environment; holding objects for robotic grasping. However, the targets of interest here are the table-top grown strawberries that naturally have different poses. Therefore, the symmetry-based idea is further developed in this study to complete 3D strawberry shapes for strawberry localisation and harvesting purposes. The contributions are summarised as follows,

- A shape completion method is proposed that aims to localise strawberry fruits more accurately. To select the best plane, a series of symmetrical planes is hypothesised and a novel rating method based on the shape of strawberries and the possible locations of mirrored points is developed. The optimal symmetrical plane that can be used to recover the complete points is used for fruit localisation.
- A test dataset is generated from reconstructed 3D strawberry data, on which the shape 20 completion methods are evaluated. Moreover, the method is implemented and evaluated on the vision system of a strawberry harvester.

2. Related work

2.1. Target localisation for fruit harvesting robot

Localisation methods can be divided into different categories according to the vision system used, including monocular vision, stereo vision and other 3D vision systems.

2.1.1. Monocular vision-based localisation

Some agricultural robots generally use a monocular vision system due to its simplicity and cost efficiency. Among these robots, some make pre-assumptions or add additional sensors to acquire 3D location. For example, Xiong et al. (2017) designed a laser weeding system with a single RGB camera for weed detection in which it was assumed that the ground is parallel to the camera frame over a fixed distance. The 3D information was calculated based on the spatial geometric relationship. Yin, Chai, Yang, and Mittal (2009) reported a tomato harvesting system that used a Charge Coupled Device (CCD) camera for tomato detection. They added a laser sensor to the system to acquire the distance information and calculated the 3D coordinates for tomato localisation. Bulanon, Okamoto, and Hata (2005) used a CCD camera for apple detection and a laser ranging sensor to measure the distance from the camera to the fruit. Their method required the target to be placed at the centre of the image and the manipulator to be controlled according to the centre when approaching the fruit.

Some monocular vision systems used localisation methods that are similar to visual servoing. De-An et al. (2011) used a CCD sensor on the end effector as a hand-eye camera for apple harvesting. The apples were localised within the image and compared with the image centre to obtain deviations. The end effector was then controlled to take small steps according to the deviation and to gradually control the arm and locate the apple. Mehta and Burks (2014) obtained the 3D position of citrus fruit using a fixed monocular camera, which assumed that the size of target citrus fruit was the average size of sample citrus fruits. Based on this assumption, the depth information could be estimated using perspective transformation. In addition to the fixed camera, the system used a hand-eye camera to regulate the end effector to locate the target fruit, leading to greater complexity and reduced speed of operation.

2.1.2. Stereo matching-based localisation

Many harvesting systems use stereo cameras as detection and localisation sensors. Font et al. (2014) proposed a fruit-harvesting system using a low-cost stereo camera to localise apples and pears. Similarly, Mehta and Burks (2016) presented a multi-camera fruit matching and localisation method using pseudo-stereo camera. The method was feasible since the fruit was sparsely dispersed in the image. Bac et al. (2014) used stereo matching to localise the stems of sweet peppers. A CCD camera was mounted on a pneumatic slide. Once the first image was taken, the pneumatic slide was moved to allow the camera to take a second image. Ji et al. (2017) used stereo matching to localise an apple branch based on skeleton points. Yang, Chang, Bao, Fan, and Xun (2018) used the stereo vision system to acquire the spatial information of White Chrysanthemums by stereo matching. Mehta, Ton, Asundi, and Burks (2017) proposed a localisation method using a stereo camera. The approach aimed at eliminating the detection errors in image processing and localising the fruit in the presence of fruit motion. However, the algorithm assumed that the stereo matching problem was solved, meaning that a given fruit must be matched in multiple cameras.

In general, localisation using stereo cameras requires the use of the matching method while the accuracy is not adequate for some precision farming purposes. Furthermore, the matching can be a problem in the outdoor environment due to the various lighting conditions.

2.1.3. Other 3D vision-based localisation

Popular 3D vision cameras include Time of Flight (TOF) camera (Gongal et al., 2015), RGB-D camera (Barnea, Mairon, & Ben-Shahar, 2016). An RGB-D camera is efficient in capturing both colour and depth images. Usually, an RGB-D camera consists of a colour sensor and two depth sensors, allowing the acquired images to not only be used in detecting the targets, but also to locate the targets in a 3D space. RGB-D cameras are also used in many harvesting and detection systems (Silwal et al., 2017; Vitzrabin & Edan, 2016; Wang et al., 2017). 3D vision-based localisation is normally straightforward when using coordinate transformation, but as mentioned above, it can only obtain a partial shape using a single camera observing a single view.

2.2. Shape completion

Shape completion in 3D is therefore essential for 3D object localisation and grasping. The following sections introduce two existing methods to obtain a complete shape. One method is to obtain the complete 3D points directly through the 3D scanning, and the other method is to complete the partial points based on shape completion.

2.2.1. Complete points from 3D reconstruction

Reconstruction methods are usually based on scanning to obtain the initial point cloud data to be used for other purposes. For example, Le Cozler et al. (2019) built a shape reconstruction system for the body of cows. Five camera–laser pairs were used to take images of a cow from different angles and locations. A complete 3D point cloud was generated by merging the point clouds from the five camera–laser pairs to monitor the growth of the cow. The entire data capturing and analysis process took around 15 min. Furthermore, similar methods were used for fruits as well. For example, Scholer and Steinhage (2015), Mack et al. (2018) utilised reconstruction methods to recreate grape bunches. They used laser range sensor or a 3D Artec spider scanner (Artec 3d, Santa Clara, CA, USA) to acquire the 3D point clouds of grape bunches for classification and segmentation. Determining complete points from 3D reconstruction usually is computationally expensive and this was obtained offline. Therefore, it is not therefore a method suited for use in real-time fruit harvesting robots.

2.2.2. *Shape completion based on partial points*

Thrun and Wegbreit (2005) proposed the idea of reconstructing a 3D partial surface using symmetry. The data being processed was partially visible surface points captured from a single 3D view. Their algorithm searched for different symmetry types and parameters and they were used to find the most plausible symmetry. The shape reconstruction could be obtained based on the proposed symmetry. Bohg et al. (2011) used a similar concept to reveal a symmetry in an incomplete view of points for the purposes of grasping. Their method assumed that a symmetric object stands on the table plane with its symmetry perpendicular to the table surface. This assumption simplifies the symmetry searching process for their specific scenario. Hypotheses were proposed based on the assumptions, and the symmetric planes were scored to find the symmetry for completing the shape.

Some researchers adopted this idea for their specific tasks. Ilonen, Bogh and Kyrki (2014) used the same assumptions and method but fused the visual and tactile sensing to achieve object reconstruction. Similarly, Schiebener et al. (2016) adapted the hypothesis testing idea, adding more hypotheses for the table planes to gain more information from the surrounding scene. Figueiredo, Moreno, and Bernardino (2017) utilised a similar method to complete 3D shapes for kitchenware objects. Makhal et al. (2018) also used the same method for shape completion, and similarly, they assumed that the symmetry was perpendicular to the table and that symmetry could be obtained via a 2D projection on the table.

The concept presented by Thrun and Wegbreit (2005) was also implemented for application to extruded shapes. For example, Kroemer, Amor, Ewerton, and Peters (2012) presented a method to complete a partial shape by using extrusions. Their method first searched for planer symmetries for extruded shapes. The detected symmetries were then used to search for suitable extrusion parameters. The proposed extrusion was rated according to a scoring system, and then the selected parameter was used to complete the extruded shape. Quispe et al. (2015) used a similar method to approximate the shape for grasping the household objects and assumed that the objects were extruded shapes. They obtained the symmetry by using the method described by Bohg et al. (2011) to propose hypotheses for the axis of extrusion. The estimated axis was optimised to improve the accuracy of the extrusion axis, which could be used for grasping.

In addition to shape completion of partial surface from symmetry, deep learning for 3D shape completion have also shown promising results (Achlioptas, Diamanti, Mitliagkas, & Guibas, 2018; Wang, Ang Jr, Lee, 2020; Yuan, Khot, Held, Mertz, & Hebert, 2018), which could be an optional method to investigate. In this paper, we mainly focus on the adoption of the concept of symmetry-based shape completion to propose a completion method for symmetric fruits, to achieve better localisation accuracy of the machine vision system in harvesting robots.

3. Materials and methods

3.1. *Procedures for the completion method*

This method was designed as a machine vision system for our table-top strawberry harvester (NORONN, www.noronn.com), as shown in Fig. 1 (a), in which the RGB-D camera senses the strawberry from the front. The algorithms of the machine vision system are briefly summarised in Fig. 1 (b). First, the detection and segmentation algorithm were designed to detect each strawberry target and segment its pixels. The dataset for training and fine tuning the model is the same as that in Ge, Xiong, and From (2019), which includes strawberry images from the Boxford Suffolk Farms (Colchester, UK) and a strawberry polytunnel at NMBU, Norway. The detection results can be found in previous publications (Ge, Xiong, Tenorio et al., 2019). The completion method was based on the segmentation results, which were further employed for 3D point extraction. The segmentation was evaluated using IoU between detected masks and the annotated ground truths. The evaluation equation is *intersection/union*, where intersection is the overlapped area and union is the area of two masks minus

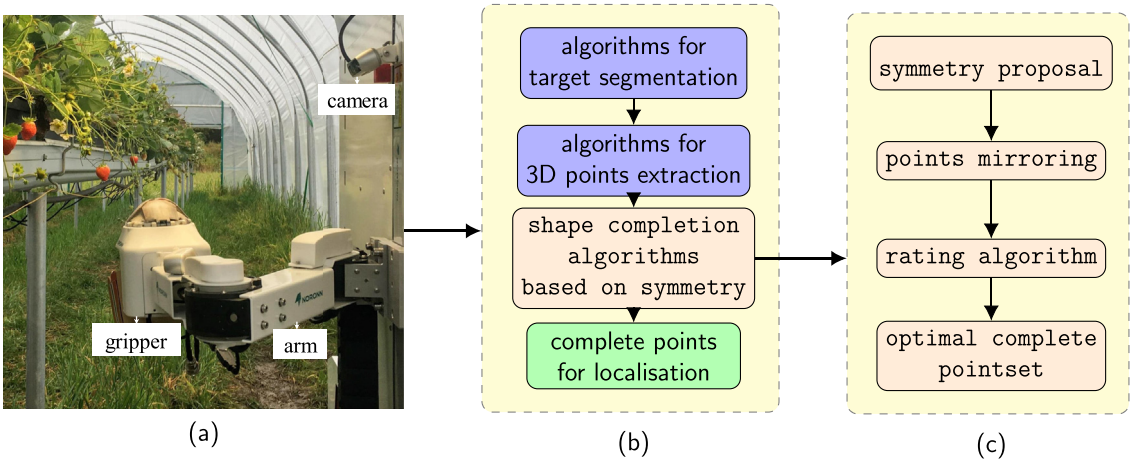

Fig. 1 – Basic concept of the machine vision system for strawberry harvesting: (a) is the harvester; (b) is the algorithm workflow of entire machine vision system; (c) is a rough workflow of the shape completion method.

the area of intersection. Sixty images were tested with 160 detected strawberries, and an average IoU of 0.86 with a standard deviation of 0.05. The result was considered accurate for the segmentation. In addition, the noise of the transformed 3D points was filtered using clustering methods (Ge, Xiong, Tenorio et al., 2019) and therefore, the extracted points were essentially good enough for the completion method. The shape completion method was subsequently designed to complete the extracted points to localise the fruit more accurately.

The basic concept of this completion method can be seen in Fig. 1 (c). The method proposes different planes of symmetry based on the position of the extracted 3D points and obtains several pointsets by mirroring the same 3D points. The core of this method is the rating algorithm, which identifies the position of each mirrored point and provides a rating based on that position. Based on the rating algorithm, the optimal pointset that represents the actual shape of the object can be obtained. The following section presents how our method can achieve shape completion based on symmetry.

3.2. Initial hypothesis

First, it is necessary to generate an initial hypothesis on which further symmetry hypotheses can be based. A 3D plane can be defined by a point $\boldsymbol{p}_0^H$ on the plane and a normal vector $\boldsymbol{n}$ that is perpendicular to the plane. The position vector of any point $\boldsymbol{p}^H$ on the plane can be expressed as:

$$\boldsymbol{n} \cdot (\boldsymbol{p}^H - \boldsymbol{p}_0^H) = 0 \tag{1}$$

Since the initial hypothesis is a rough estimate of the plane used for generating more hypotheses, average coordinates of all the points in the point cloud P are used as the point on the initial plane. A vector that is parallel to the z axis of the camera coordinate frame was used to represent the normal vector of the initial hypothetical plane.

3.3. Generation of symmetric plane hypotheses

The generation of hypothetical planes includes translation and rotation of the initial hypothetical plane. The translation occurred along the depth direction over a regular distance interval Δd and the rotation was along the x and y axis over regular angle intervals $\Delta\theta_1$ and $\Delta\theta_2$. To perform the translation, the point defined on the initial plane was moved a specified distance along the depth direction. $R \cdot \boldsymbol{n}$ was used to perform a local rotation and generate the normal vector of new hypothesis,

$$\boldsymbol{n}_{new} = R \cdot \boldsymbol{n}_0 \tag{2}$$

where R is the rotation matrix generated by $\Delta\theta_1$ and $\Delta\theta_2$ as shown in Eq. (3).

$$R = \begin{bmatrix} \cos\Delta\theta & -\sin\Delta\theta & 0 \\ \sin\Delta\theta & \cos\Delta\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

In this way, $m \times \Delta d$, $j \times \Delta\theta_1$ and $k \times \Delta\theta_2$ can be used to generate a total of $m \times j \times k$ hypotheses, in which m, j and k represent the number of translations along depth, the number of rotations along x, the number of rotations along y, respectively. The process is illustrated in Fig. 2. In our algorithm for strawberry, m, j and k was set to 4, 11 and 3, respectively. Less rotations along y axis were generated because a strawberry can be mostly regarded as axisymmetric body along y axis. Δd, $\Delta\theta_1$ and $\Delta\theta_2$ were set to $0.2 \times$ *len depth*, 5°, and 5°, respectively, where *len depth* is the length of depth of detected strawberry. The values of the number of hypothesis could be larger while the values of interval length could be smaller, so that more situations can be included. However, by doing so, more hypothetical planes would be generated, which would make the algorithm slower. In our algorithm, these values were chosen based on the size of strawberries and the need to balance speed and performance.

3.4. Points mirroring

To obtain the complete shape the visible points need to be mirrored based on the hypothetical plane. Furthermore, the mirrored points will be used for rating during the next section. In order to mirror the original point cloud P through the proposed symmetric plane, the line equation in 3D space was utilised. Figure 3 illustrates the calculation progress in 3D space. The points can be regarded as vectors, marked by yellow dotted lines in Fig. 3. A line in 3D space can be determined by a point on the line and a vector parallel to the line. Thus, a line $\boldsymbol{p}^L$ that passes through the point $\boldsymbol{p}_0^L$ in the original point

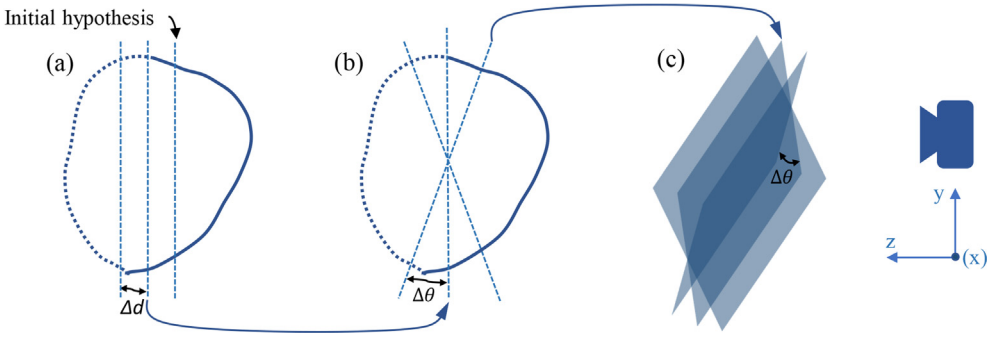

Fig. 2 – Generation progress of the plane hypotheses: (a) shows the initial hypothetical plane and its two translation planes; (b) shows the rotated planes along x axis; (c) shows rotated planes in 3D space along y axis; axes on the right shows the perspective of the camera.

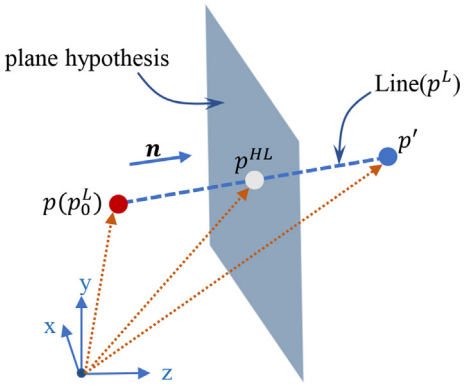

Fig. 3 – Schematic for points mirroring. The plane in the middle is the hypothetical plane; $\boldsymbol{p}$ represents a point on the strawberry while it is also a point $\boldsymbol{p}_0^L$ on the line; $\boldsymbol{p}'$ represents the mirrored point.

cloud P and is perpendicular to the plane can be defined as follows:

$$\boldsymbol{p}^L = c^*\boldsymbol{n} + \boldsymbol{p}_0^L \tag{4}$$

where $\boldsymbol{n}$ is the normal vector of the plane and c is a constant to determine the point position on the line. This line also passes through the point $\boldsymbol{p}'$ in the mirrored point cloud P' and the point $\boldsymbol{p}^{HL}$, which represents the intersection point between the line and the plane hypothesis. Because $\boldsymbol{p}^{HL}$ also satisfies the plane equation Eq. (1), c in Eq. (4) and the coordinates of $\boldsymbol{p}^{HL}$ can be obtained by replacing $\boldsymbol{p}^L$ with $\boldsymbol{p}^{HL}$ in the Eq. (4). Since the intersection point is the midpoint of $\boldsymbol{p}$ and $\boldsymbol{p}'$, the mirrored point can be obtained as follows:

$$\boldsymbol{p}' = 2\boldsymbol{p}^{HL} - \boldsymbol{p} \tag{5}$$

where of $\boldsymbol{p}$, $\boldsymbol{p}^{HL}$ and $\boldsymbol{p}'$ can be regarded as vectors from the coordinate origin to the corresponding scatter point, as illustrated by the orange lines in Fig. 3.

3.5. *Rating for generated hypotheses*

3.5.1. *Locations definition*

The next step is to rate the mirrored points to find the optimal hypothesis that can be used to obtain the complete strawberry shape. As shown in Fig. 4 (a), a point p in the mirrored point cloud P' has the possibility to be located in four places. If a mirrored point lies in locations 1 and 2 in Fig. 4 (a), it should already been sensed as one of the points in the original point cloud. Therefore, the mirrored point of locations 1 and 2 in Fig. 4 (a) decreases the possibility of the corresponding point cloud being a good candidate. If a mirrored point lies in location 3 in Fig. 4 (a), it supports the hypothesis. If a mirrored point lies in location 4 in Fig. 4 (a), it may be a supported point that indicates the unseen side of the object. Thus, locations 3 and 4 in Fig. 4 (a) were considered as positive cases, and 1 and 2 as negative cases.

However, location 4 in Fig. 4 (a) covers a large part of the area, as can be seen in the region marked with red dashed lines. If the rating for location 4 in Fig. 4 (a) purely depends on the distance to the nearest point in the original points, as in the method described by Bohg et al. (2011), the mirrored points may gather behind the visible points, resulting in inaccurate shape recovery. Therefore, the area of location 4 in Fig. 4 (a) was divided into three additional cases as shown in Fig. 4 (b). Locations 4–1 and 4–3 in Fig. 4 (b) are the negative cases since they are outside the reasonable region, while location 4–2 in Fig. 4 (b) is the positive case because it is within the reasonable region. The reasonable region is marked with red dashed lines in Fig. 4 (b).

3.5.2. *Location identification*

Before identifying where a point is located, some pre-processing and distance calculations are required. The parameters defined in this section can be seen in Fig. 5. The original points P and mirrored points P' were projected to the image plane to generate the corresponding mask Mp and Mp'. Three distances were calculated and utilised for rating every point p' in P'. The first is a reference distance $Dist_ref$ of the size

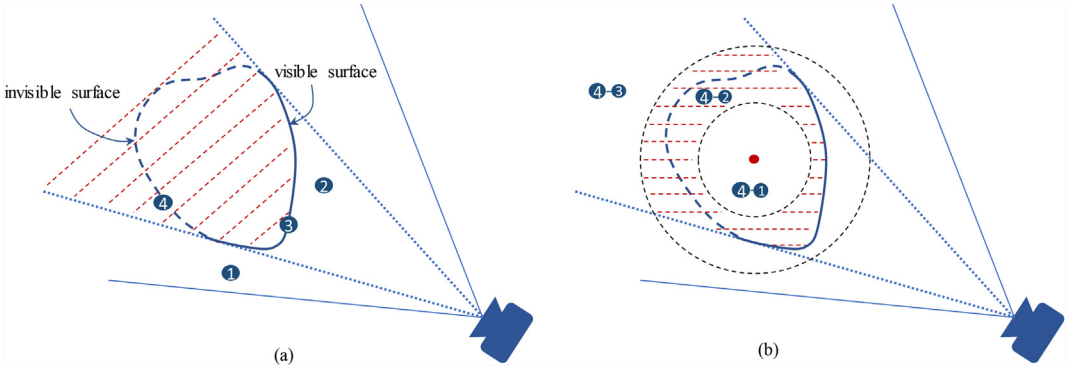

Fig. 4 – Different location cases for the mirrored points: (a)-location 1: the mirrored point is outside the detected mask; (a)-location 2: the mirrored point is in front of the original point cloud; (a)-location 3: the mirrored point is exceptionally close to the original points; (a)-location 4: the mirrored point is in the invisible area behind the detected mask, marked with red dashed lines; (b)-location 4–1:the mirrored points are close to the centre of the object; (b)-location 4–2: the mirrored points are in the reasonable region; (b)-location 4–3: the mirrored points are far from the reasonable region.

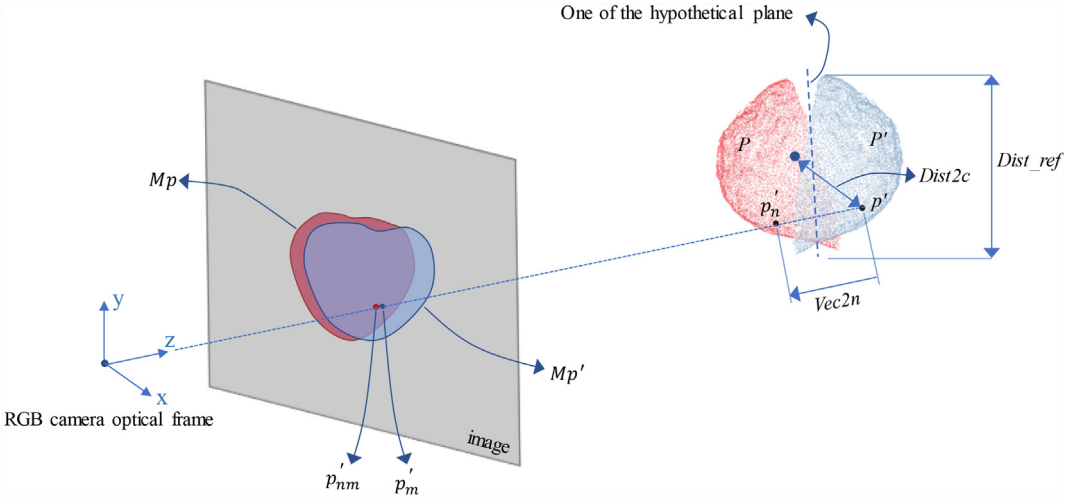

Fig. 5 – Illustration of location identification: the axes on the left represent the axes of the RGB camera optical frame; the red points on the right are the original points of strawberry while the blue points are the corresponding mirrored points through the hypothetical plane; the plane in the middle shows the projected masks of the original points and mirrored points on the image plane.

of the target. The most substantial distance among the three directions x, y and z was taken as the reference distance *Dist_ref*. The distance of every point p' in P′ was calculated in relation to the centre of original points *Dist2c* and the depth vector *Vec2n* from p' to p'_n. Points p'_{nm} and p'_m represent the projected points of p'_n and p' in the projected mask *Mp* and *Mp′*. The procedure aims to find the nearest point of p'_n in *Mp*, which is p'_{nm}, which was then de-projected to the 3D space to obtain p'_n, so that the distance between p' and p'_n could be calculated.

The identification of the location of a point is described in this section. For a point p' in P′, if the projected point in *Mp′* of point p' in P′ is outside mask *Mp*, then it is recognised as location 1. Otherwise, it is recognised as one of the other three locations. For the remaining points, if the distance ***Vec2n*** is negative and smaller than a predefined value *V2n_pre* around zero, then it is recognised as location 2. If the distance ***Vec2n*** is close to zero, then it is recognised as location 3. The remaining points belong to location 4 and are divided into three more locations. If *Dist2c* is smaller than the predefined distance *D2c_pre*, it denotes location 4–1, while if *Dist2c* is larger than the predefined distance *D2c_pre*, then it is considered as being the location 4–3. Otherwise, it signifies location 4–2.

3.5.3. *Rating algorithm*

The rating metrics for different locations are shown below. When the point is at location 1, the score is negative. The larger the distance between the point and its nearest point ***Vec2n***, the lower the score. The equation can be defined as:

$$Score(p') = -k_1 * \frac{|\boldsymbol{Vec2n}|}{Dist_ref} \quad (6)$$

where, k_1 is a predefined constant used to adjust the rating. The study set $k_1 = 10$ so that if $|\boldsymbol{Vec2n}|/Dist_ref$ is between 0 and 0.1, the score is between −1 and 0.

If the point is at location 2, the score is negative, as shown in Eq. (7). If the point is further away from the nearest point, the score is lower. A tolerant distance $1/k_2$ was set to define location 2. If the ***Vec2n*** is less than $-\frac{1}{k_2} * Dist_ref$, it was recognised as location 2 and scored using the equation below:

$$Score(p') = -k_2 * \frac{|\boldsymbol{Vec2n}|}{Dist_ref} \quad (7)$$

The score of location 3 is positive because it is a positive case, which can be expressed as:

$$Score(p') = k_3 \quad (8)$$

where, a constant number $k_3 = 1$ was used to score this case.

The score of location 4–1 is negative, as shown in Eq. (9). If the point is closer to the centre point, the score is lower

$$Score(p') = -k_4 * \frac{Dist_ref}{Dist2c} \quad (9)$$

Also, the score of location 4–3 is negative, as shown in Eq. (10). If the point is further away from the centre point, the

Table 1 – List of parameters for location identification and rating.

Usage	No.	Parameters	Values
Location identification	1	*Dist_ref*	calculated
	2	*Dist2c*	calculated
	3	***Vec2n***	calculated
	4	*Dist_tol*	j_1*Dist_ref, $j1 = 0.05$
	5	*Dist2c_0*	j_2*Dist_ref, $j2 = 0.6$
	6	*Dist2c_1*	j_3*Dist_ref, $j3 = 1.2$
Rating	7	k_1	10
	8	k_2	1/j1
	9	k_3	1
	10	k_4	1/j2
	11	k_5	1/j3

Algorithm 1
Rating for the proposed candidates of the mirrored points.

Data: $P' = \{p' \in R^3\}$
$Mp = \{p_M \in R^2\}$
$Mp' = \{p'_M \in R^2\}$
Result: $Score(P')$
for every $p' \in P'$ **do**
  comparing the strawberry position with strap line and table surface;
  if Mp' outside Mp **then**
    Calculate the score s^i using Eq. (6)
  else if ***Vec2n*** is opposite to z axis direction and $|Vec2n|$ is larger than $Dist_tol$ **then**
    Calculate the score s^i using Eq. (7)
  else if $|Vec2n|$ is smaller than $Dist_tol$, **then**
    Calculate the score s^i using Eq. (8)
  else if $Dist2c$ is smaller than $Dist2c_0$ **then**
    Calculate the score s^i using Eq. (9)
  else if $Dist2c$ is larger than $Dist2c_1$ **then**
    Calculate the score s^i using Eq. (10)
  else
    Calculate the score s_i using rating Eq. (8)
end
Calculate an overall $Score(P')$ for the points set by averaging all scores

score is lower. The score of location 4–2 is a positive case, and its equation is same as Eq. (9).

$$Score(p') = -k_5 * \frac{Dist2C}{Dist_ref} \tag{10}$$

The parameters appearing in section 3.5 are listed in Table 1. The parameters 1–6 were used for the location identification of mirrored points. *Dist_ref*, *Dist2c* and ***Vec2n*** are calculated based on the size of detected strawberry, the distance between mirrored point and the detected strawberry centre and vector from mirrored point to nearest point, respectively, as can be seen in Fig. 5. *Dist_tol*, *Dist2c_0* and *Dist2c_1*are used in location identification, by comparing with *Dist2c* and ***Vec2n***, their values are shown in Table 1, in which *Dist_tol* is a tolerant distance to define location 2 and it should be a thin region along the strawberry surface, so j_1 is set to 0.05. *Dist2c_0* and *Dist2c_1* are the radii of two circles, used to set the boundary of location 4–2, as shown in Fig. 4. The values of *Dist2c_0* and *Dist2c_1* were set to 0.6 and 1.2 of *Dist_ref* and the values were defined according to the size and shape of strawberry and a few trials during the test.

Parameters k_1 to k_5 were used as rating adjustment constants, as can be seen from Eq. (6) to Eq. (10). The criteria for defining these values is to balance the impact of each score on the overall score. The score of a mirrored point was set to 1 if it is within the reasonable region, while its score was set in the range of (−1, 0) if it is not in the reasonable region but close to it. In addition, the score of a point was set to less than −1 if it is considered to be far from the reasonable region. Based on the above criteria, location 3 and location 4–2 were positive cases within the reasonable region, then the score for them was set to the same value k_3, which is 1. For location 2, k_1 was set to 10 so that when $|Vec2n|/Dist_ref$ is between 0 and 0.1, the score is in the range of (−1, 0). k_2 was set to $1/j1$, so that the score of location 2 is in the range of (−1, 0), while k_4 was set to $1/j2$, so that the rating score of location 4–1 is in the range of (−1, 0). The value of k_5 was set to 1 in this study so that the score would be smaller than −1.

The rating algorithm for one pointset is shown in Algorithm 1. The outlined rating flow was applied to every generated candidate of the mirrored points, to obtain the scores $\{S_1, S_2, \ldots, S_n\}$. for a the candidates $P'_1, P'_2, \ldots, P'_n$. The best candidate is represented by the mirrored points with the highest average score $S_{best} = max\{S_1, S_2, \ldots, S_n\}$. The first pointset appeared in the array was determined as the optimal candidate if several pointsets were scored the same.

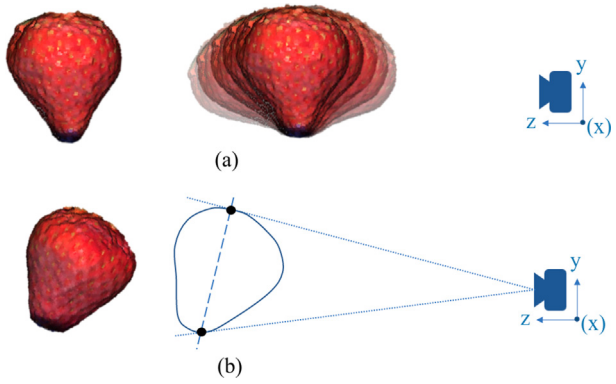

Fig. 6 – Visible points generation: (a) is to generate different poses of strawberry; (b) is to generate visible points from the complete point cloud.

3.6. *Estimation of the 3D position*

Based on the proposed shape completion method, a 3D location of the target strawberry can be estimated. The range of all the strawberry points was extracted in three coordinate directions to build up a cube. The position of this cube was used to represent the strawberry location in the 3D space and was sent to the robot control system.

4. Evaluation on the reconstructed point cloud

4.1. *Data preparation from the reconstructed data*

The first evaluation was performed on the fully reconstructed 3D strawberry data, which are dense point clouds from 3D scanning of real strawberries. In this case, the ground truth of the 3D position of the strawberry could be accurately defined. The data was kindly provided by He, Harrison, and Li (2017), who implemented a reconstruction method to obtain point clouds from various strawberries. The reconstructed point cloud is dense with more than ten thousand points, which is unnecessary in a real-time robotic system and, in turn, increase the processing time. To speed up the algorithm, a down-sampling process was conducted. Six hundred points were randomly selected from the reconstructed point cloud, which represents the approximate number of points for one strawberry from the RGB-D camera, in terms of the machine vision system in our strawberry harvester.

In a practical table-top scenario, strawberries have different poses, this may affect the results of shape completion. Therefore, the down-sampled points were rotated along the x to generate targets with different poses. The initial position was set when the axis of strawberry is approximately parallel to the y axis of the camera. The rotation was

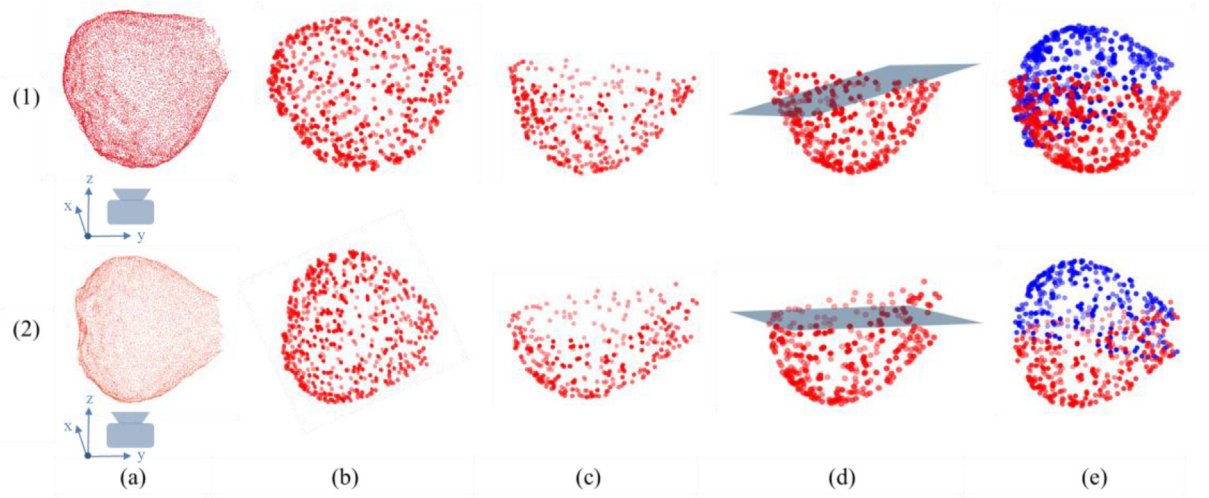

Fig. 7 – Two examples regarding the visualisation of the results of the reconstructed strawberry data: (a). the reconstructed dense strawberry point cloud; (b). down-sampled points; (c). the invisible points removed; (d). the visible points with the best symmetry hypothesis; (e). the completed strawberry points, including visible points and mirrored points.

performed every five degrees, up to 45° degrees, in a clockwise and anti-clockwise direction. The generation process is roughly illustrated in Fig. 6 (a). During this process, 19 poses were generated for each strawberry point cloud, and the data from ten reconstructed strawberry point clouds were used. Therefore, the instances to be tested were 190 in total.

After the rotation, the invisible points were removed from the generated points allowing the remaining points to represent the visible points from the camera. The method of generating visible points is shown in Fig. 6 (b). The first step is to find the two points with the smallest and largest y coordinates in the visible area of the camera. These two points form a plane that is also perpendicular to the *xz* plane. The points in front of the plane are visible to the camera. In this way, the visible 3D points were generated and could be used as input for the shape completion algorithm.

4.2. *Evaluation and results*

Figure 7 shows two examples of the data generation processes (Fig. 7 (b) and (c)) and the shape completion results (Fig. 7 (d) and (e)). The coordinate system markers in the image show how the camera perceived the strawberry, where z is the depth direction. Figure 7 (1) signifies a fruit whose larger side with calyx can be seen from the camera, while Fig. 7 (2) presents one where the tip side that can be seen from the camera. The results in Fig. 7 (e) show that the method can complete the shape in the correct direction.

The shape completion aimed to estimate the location of each fruit accurately. The 3D bounding box position was sent from the machine vision system to the robot control system (Xiong, Ge, Grimstad, & From, 2020). Therefore, the bounding box was utilised to determine if the shape completion can improve the accuracy of fruit localisation. The IoU between the generated complete points and the ground truth points was calculated, as well as the IoU between the visible input points and the ground truth points. The equation can be expressed as follows:

$$\mathrm{IoU} = \frac{ovelapping\ volumn}{combined\ volumn} \tag{12}$$

where the overlapping volume signifies the volume of the overlapping cube shown in Fig. 8.

The combined volume is the "ground truth" volume plus the "generated volume" minus the "overlapping" volume. In addition to the IoU in 3D, the deviation of the centre of the bounding box between the ground truth and visible points as well as the completed shape, was also calculated. These two measurements are presented in Fig. 8. So far, the data and metrics for evaluation have been defined, and the evaluation procedures are now outlined in Algorithm 2.

Figures 9 and 10 show the evaluation results from all the different angles provided by the strawberries data, from −45° to +45°, using box plots. The cross within the box denotes the mean, while the short line within the box signifies the median. The length of the box reflects the magnitude of the deviation, while the maximum and minimum values can also be found on the boxes. Therefore, the IoUs of the completed points are larger than those of the visible points while the centre deviations of the completed points are smaller than those of the visible points.

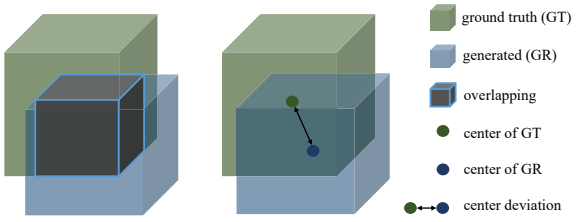

Fig. 8 – IoU and centre deviation in 3D.

Algorithm 2
Rating procedure for proposed candidates of mirrored points.

Data: $P^r = \{p_i^r \in R^3\}$
Result: all IoUs and centre deviations
for every constructed pointset P^r **do**
  Down-sample the dense point cloud P^r to obtain a point set P^s with 600 points
  Save the 3D position and centre of randomly chosen points
  Rotate the selected points P_k^s along the x axis to generate $2 * k + 1$ of new pointset $P_1^s, P_2^s, \ldots, P_{2k+1}^s$, k is the defined number of pointset on one side of the initial position.
  for every generated pointset P^s **do**
    Generate the visible points P
    Input P to the algorithm
    Save the 3D position and the centre of visible and completed points
    Calculate intersection of unions (IoUs), centre deviations
  end
end

The average value of IoU and centre deviation are shown in Table 2. The IoU between the results and ground truth is 0.77, and the centre deviation is 6.9 mm, while the IoU and centre deviation of unprocessed partial points are 0.56 and 14.1 mm.

5. Application in real environment

5.1. *System architecture of the machine vision system*

Various methods have been used by previous research work to obtain the 3D target position for a harvesting robot. The method presented in this paper used Mask R–CNN process (He and Gkioxari et al., 2017) for the instance segmentation and acquired 3D positions through coordinate transformation (Ge, Xiong, Tenorio et al., 2019). The system architecture of the detection, segmentation and shape completion is presented in Fig. 11. As illustrated in Fig. 11, the 3D points of the target fruits were obtained via camera observation and perception to facilitate further shape completion.

5.2. *Data from the RGB-D camera*

An RGB-D camera (D435, Intel, USA) was employed to capture data, and the subsequent point cloud that is presented in Fig. 12 (a), which shows that only the front side of these strawberries is visible to the camera. Therefore, the shape completion method was used here to estimate the complete position of these fruits.

Figure 12 (b) shows four examples of strawberries that were extracted from the visible side of the RGB-D camera. Furthermore, 30 strawberry instances were extracted from the point cloud and the ground truth of bounding boxes were manfully annotated.

5.3. *Evaluation and results*

Figure 13 shows the two examples of visualised input points from an RGB-D camera and the corresponding completion results. Figures 13 (a), (b) and (c) show the extracted 3D points, the points with the optimal hypothesis plane and the completed points, respectively. In addition, Fig. 14 shows three examples of the detection and shape completion results visualised in point cloud with bounding boxes. The segmented masks of strawberries using Mask R–CNN, as shown in Fig. 14 (a), were transformed from the image frame to the RGB camera optical frame, as signified by the black points in Fig. 14 (d) and (e), while the white points in Fig. 14 (e) represent the points of the other completed part. Fig. 14 (b) and (c) denote the original point cloud from camera, in which the original strawberry points are visible. The bounding boxes in Fig. 14 (b)–(e) enclose the whole competed points.

A test was conducted to evaluate the time requirements of the completion method. A total of 132 cases were used for the testing. With the current number of hypotheses, the average time to execute the completion was 0.26 s with a standard deviation of 0.01s. Our harvesting system uses static picking (Xiong, Peng, Grimstad, From, & Isler, 2019, 2020), which means that the platform stops at certain point until and it finishes the picking in current image frame. The machine vision system excluding the completion method take an average of 0.82 s per image frame (Ge, Xiong, Tenorio et al., 2019). Therefore, the processing time of one to two seconds for one image frame is acceptable to our system.

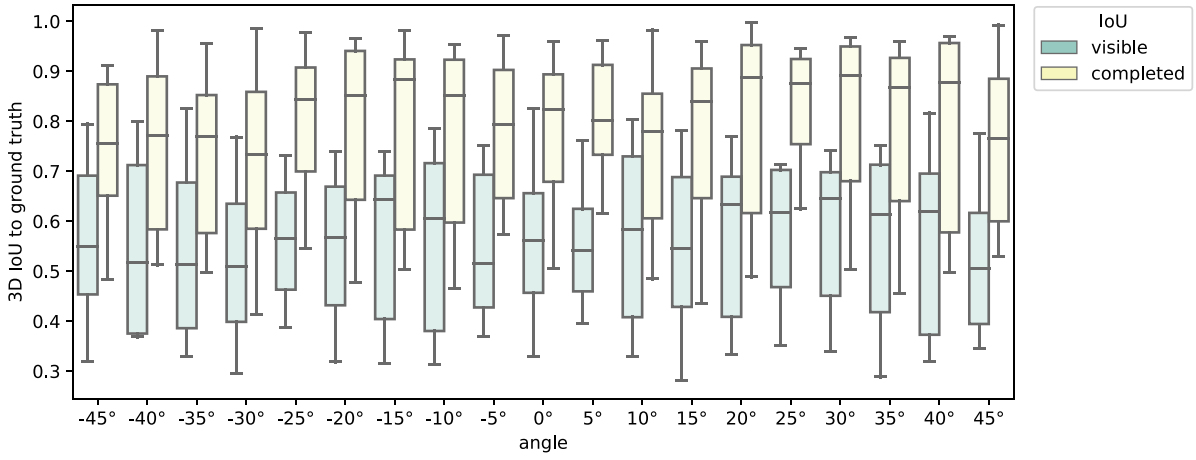

Fig. 9 – The overall evaluation results of centre deviations according to the data derived from the 3D reconstruction.

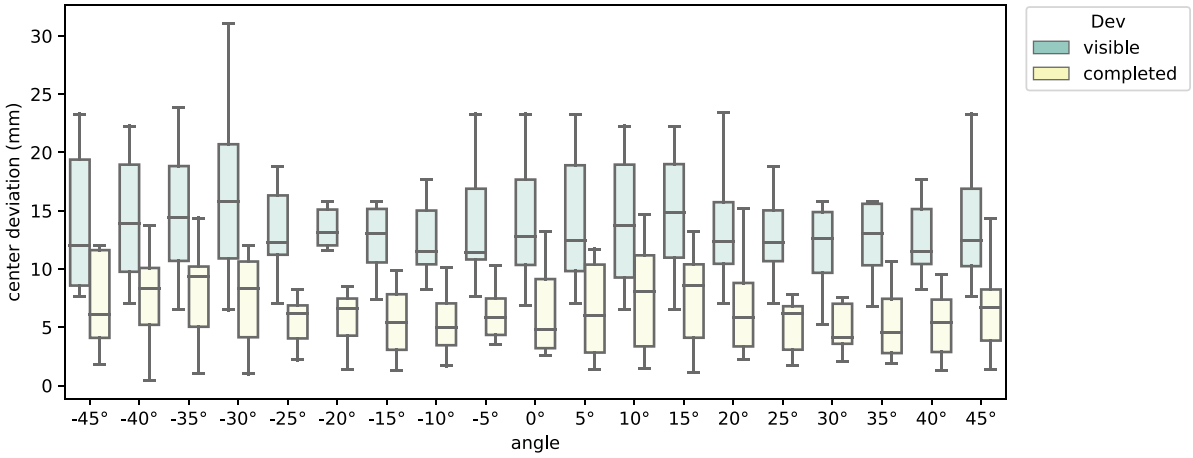

Fig. 10 – The overall evaluation results of centre deviations according to the data derived from the 3D reconstruction.

Table 2 – The average evaluation results of pointsets of strawberry derived from the 3D reconstruction.

Average IoU		Average Centre Dev	
Input points	Generated points	Input points	Generated points
0.56	0.77	14.1 mm	6.9 mm

The IoU of the 3D bounding boxes and the centre deviation were established by comparing the visible points and completed points with the ground truth. The ground truths were manually annotated for all targets. The overall results are shown in Fig. 15, including mean, deviation, median, maximum and minimum values. The average results are listed in Table 3. The IoU between the generated points and ground truth is 0.61, and the centre deviation is 5.7 mm, while the IoU and centre deviation of unprocessed partial points are 0.47 and 8.9 mm, respectively.

5.4. *Discussions and limitations*

The completion algorithm was performed after the strawberry points have been extracted. The pre-processing steps include strawberry detection, strawberry pixel segmentation and coordinate transformation based on the segmented pixels. The pixel segmentation and coordinate transformation have been tested robust enough during the experiments. Detection could be a problem when an unripe strawberry was detected as a ripe strawberry, resulting a failure picking. However, we believe that this problem does not influence the completion algorithm.

When the point clouds from the camera are accurate, accurate 3D strawberry points could be obtained through coordinate transformation and the clustering method. However, the strawberry points might be dragged to the front and back along the depth direction by adjacent objects, such as leaves, other strawberries, or stems. This results in deformed points. It can be seen that the results obtained from the 3D

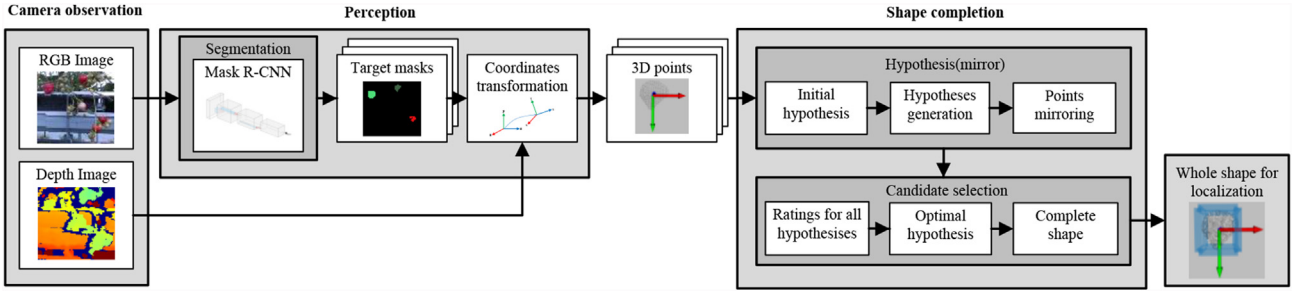

Fig. 11 – System architecture for shape completion in a real environment: the RGB image of the fruits is segmented using Mask R–CNN trained on corresponding fruit data; each of the resulted masks represents a target fruit whose coordinates are transformed from the image frame to the RGB camera optical frame; then the 3D points of the target fruits are obtained, which can be further processed by the shape completion method.

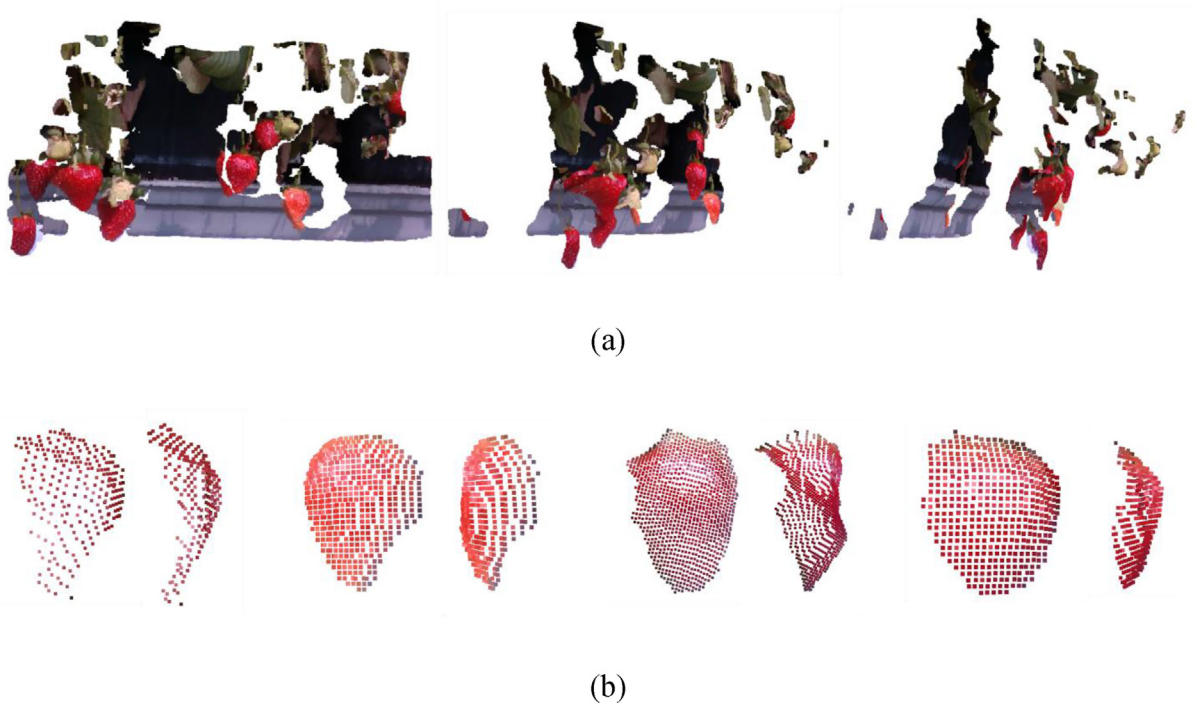

Fig. 12 – Point cloud data: (a) the visualisation of the point cloud data of the table-top strawberries: from the left to right is normal view (front view), side-front view and side view respectively; (b) four examples of data extracted from the point cloud: for each example, left is the front view, right is the side view.

reconstructed data were superior to the data provided by RGB-D camera, indicating that the reconstructed data are more accurate, complete, and can represent the actual surface of the target, while the data from the RGB-D camera may be deformed due to the limited accuracy of the depth camera. Therefore, the completion method requires further improvement by advancing the performance of camera and refining of surface points.

Occlusions can cause more severe deformation because it is difficult for the camera to perfectly transition between

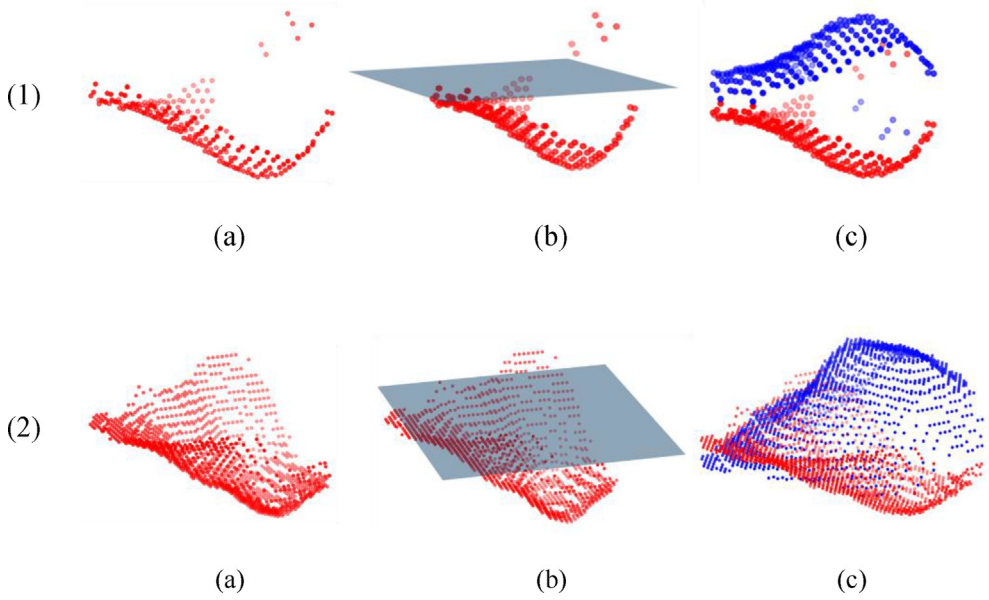

Fig. 13 – Two examples of visualisation of results according to the reconstructed strawberry data: (a) extracted 3D points; (b) visible points with the best hypothesis of symmetry; (c) completed strawberry points, including visible points and mirrored points.

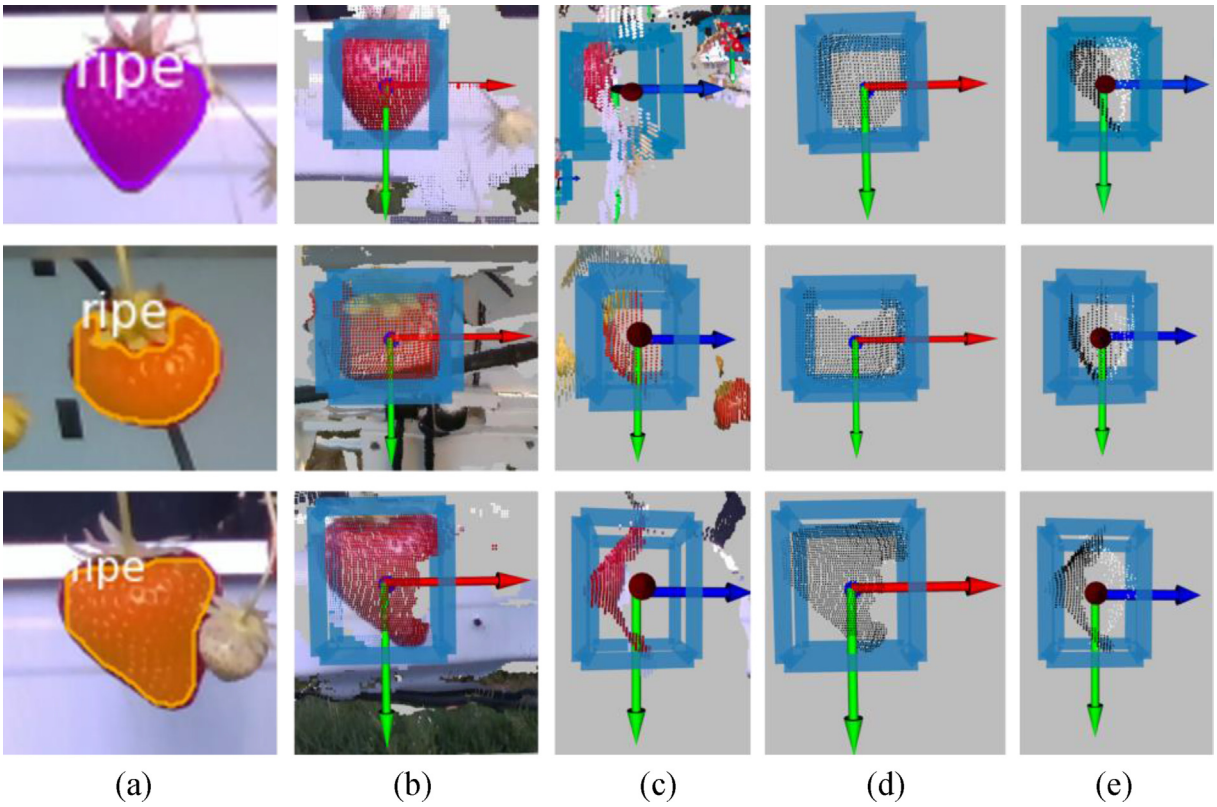

Fig. 14 – The visualisation of the shape completion results, in which the red, green and blue axes represents *x*, *y* and *z* directions, respectively and the bounding boxes enclose the completed strawberry points: (a) represents the segmentation results of strawberries using Mask R–CNN; (b) and (c) represent the front view and side view of the original strawberry point cloud; (d) and (e) represent the front view and side view of the completed points, including the extracted strawberry point cloud obtained via segmentation and coordinate transformation, and represented by the black points, as well as the point cloud of the other half part obtaining from the shape completion, and represented by the white colour.

objects, therefore may influence the results of completion. Slight occlusions were included in the experiments, while severe deformations caused by severe occlusions were not included because severely deformed strawberry point clouds provide inaccurate location information and therefore, lose their role in localisation. In addition to using a camera with better accuracy, future work could investigate the algorithm to identify severely deformed points so that they can be skipped in the current image view.

Another factor that affect the results is the strawberry position pose. As introduced in section 3.1 and can be seen in Fig. 1 (a), our camera has a front view of strawberry tabletops. The pose of the strawberries, as seen from the perspective of the camera, is mostly visible along the long axis of the fruit. Therefore, the method is only effective in this type of system setting, where the strawberries are observed along their long axis.

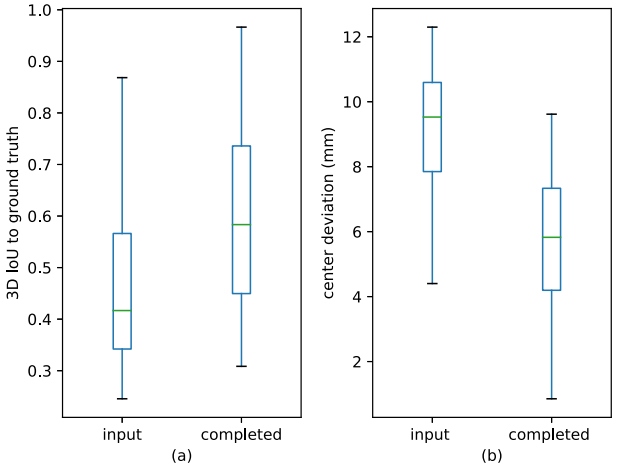

Fig. 15 – The overall evaluation results of the data from the RGB-D camera: (a) represents the evaluation results of 3D IoU on the data from the RGB-D camera; (b) represents the evaluation results of the centre deviation on the data from RGB-D camera.

Table 3 – The average evaluation results of the data from the RGB-D camera.

Average IoU		Average Centre Dev	
Input points	Generated points	Input points	Generated points
0.47	0.61	8.9 mm	5.7 mm

6. Conclusions

This paper proposed a shape completion method for a strawberry-harvesting robot for accurate fruit localisation. Based on the observation that strawberries are symmetrical in shape, several hypothetical planes were proposed as symmetries for shape completion. For every proposed hypothesis for the symmetric plane, the corresponding mirrored pointset can be obtained from the visible points. A rating system was proposed to identify the optimal hypothetical plane of symmetry and the corresponding mirrored pointset, based on the positions of the mirrored points. The entire shape of the strawberry fruit can be obtained by merging the mirrored points and the visible points.

The method was evaluated on completely reconstructed strawberry point clouds and implemented in a strawberry harvesting robot using an RGB-D camera in a farm setting where table-top strawberries were grown. The evaluation of 3D IoU showed that the 3D bounding boxes of the completed points could align with the ground truth by 0.77 and 0.61 for the reconstructed point cloud and the data from the RGB-D camera, respectively, while the unprocessed points align with the ground truth by 0.56 and 0.47 for the reconstructed point cloud and the data from the RGB-D camera, respectively. Also, using this method, the centre deviations between the completed points and the ground truth were 6.9 mm and 5.7 mm for the test data generated from reconstructed point clouds and from the RGB-D camera, respectively, while the ones found between unprocessed points and the ground truth were 14.1 mm and 8.9 mm, for the test data generated from reconstructed point clouds and from the RGB-D camera, respectively. Therefore, the proposed method presented accurate localisation results for the target fruits of the strawberry-harvesting robot. Furthermore, we think the proposed completion method could also be applicable to other fruits with symmetries, such as apples and oranges.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by The Research Council of Norway [grant number 2962020].

REFERENCES

Achlioptas, P., Diamanti, O., Mitliagkas, I., & Guibas, L. (2018, July). Learning representations and generative models for 3d point clouds. *International conference on machine learning* (pp. 40–49).

Bac, C. W., Hemming, J., & Van Henten, E. J. (2014). Stem localization of sweet-pepper plants using the support wire as a visual cue. *Computers and Electronics in Agriculture, 105*, 111–120.

Bac, C. W., Hemming, J., Van Tuijl, B. A. J., Barth, R., Wais, E., & van Henten, E. J. (2017). Performance evaluation of a harvesting robot for sweet pepper. *Journal of Field Robotics, 34*(6), 1123–1139.

Barnea, E., Mairon, R., & Ben-Shahar, O. (2016). Colour-agnostic shape-based 3D fruit detection for crop harvesting robots. *Biosystems Engineering, 146*, 57–70.

Bohg, J., Johnson-Roberson, M., León, B., Felip, J., Gratal, X., Bergström, N., … Morales, A. (2011, May). Mind the gap-robotic grasping under incomplete observation. *2011 IEEE International Conference on Robotics and Automation* (pp. 686–693). IEEE.

Bulanon, D. M., Okamoto, H., & Hata, S. I. (2005). Feedback control of manipulator using machine vision for robotic apple harvesting. *2005 ASAE Annual Meeting* (p. 1). American Society of Agricultural and Biological Engineers.

De-An, Z., Jidong, L., Wei, J., Ying, Z., & Yu, C. (2011). Design and control of an apple harvesting robot. *Biosystems Engineering, 110*(2), 112–122.

Figueiredo, R., Moreno, P., & Bernardino, A. (2017, February). Automatic object shape completion from 3d point clouds for object manipulation. In *International joint conference on computer vision, imaging and computer graphics theory and applications, 4* pp. 565–570).

Font, D., Pallejà, T., Tresanchez, M., Runcan, D., Moreno, J., Martínez, D., … Palacín, J. (2014). A proposal for automatic fruit harvesting by combining a low cost stereovision camera and a robotic arm. *Sensors, 14*(7), 11557–11579.

Ge, Y., Xiong, Y., & From, P. J. (2019). Instance segmentation and localization of strawberries in farm conditions for automatic fruit harvesting. *IFAC-PapersOnLine, 52*(30), 294–299.

Ge, Y., Xiong, Y., Tenorio, G. L., & From, P. J. (2019). Fruit localization and environment perception for strawberry harvesting robots. *IEEE Access, 7*, 147642–147652.

Gongal, A., Amatya, S., Karkee, M., Zhang, Q., & Lewis, K. (2015). Sensors and systems for fruit detection and localization: A review. *Computers and Electronics in Agriculture, 116*, 8–19.

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *Proceedings of the IEEE international conference on computer vision* (pp. 2961–2969).

He, J. Q., Harrison, R. J., & Li, B. (2017). A novel 3D imaging system for strawberry phenotyping. *Plant Methods, 13*(1), 1–8.

Ilonen, J., Bohg, J., & Kyrki, V. (2014). Three-dimensional object reconstruction of symmetric objects by fusing visual and tactile sensing. *The International Journal of Robotics Research, 33*(2), 321–341.

Ji, W., Meng, X., Qian, Z., Xu, B., & Zhao, D. (2017). Branch localization method based on the skeleton feature extraction and stereo matching for apple harvesting robot. *International Journal of Advanced Robotic Systems, 14*(3), 1729881417705276.

Kroemer, O., Amor, H. B., Ewerton, M., & Peters, J. (2012). November). Point cloud completion using extrusions. *2012 12th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2012)* (pp. 680–685). IEEE.

Le Cozler, Y., Allain, C., Caillot, A., Delouard, J. M., Delattre, L., Luginbuhl, T., et al. (2019). High-precision scanning system for complete 3D cow body shape imaging and analysis of morphological traits. *Computers and Electronics in Agriculture, 157*, 447–453.

Lehnert, C., English, A., McCool, C., Tow, A. W., & Perez, T. (2017). Autonomous sweet pepper harvesting for protected cropping systems. *IEEE Robotics and Automation Letters, 2*(2), 872–879.

Mack, J., Lenz, C., Teutrine, J., & Steinhage, V. (2017). High-precision 3D detection and reconstruction of grapes from laser range data for efficient phenotyping based on

supervised learning. *Computers and Electronics in Agriculture, 135*, 300–311.
Mack, J., Schindler, F., Rist, F., Herzog, K., Töpfer, R., & Steinhage, V. (2018). Semantic labeling and reconstruction of grape bunches from 3D range data using a new RGB-D feature descriptor. *Computers and Electronics in Agriculture, 155*, 96–102.
Makhal, A., Thomas, F., & Gracia, A. P. (2018, January). Grasping unknown objects in clutter by superquadric representation. *2018 Second IEEE International Conference on Robotic Computing (IRC)* (pp. 292–299). IEEE.
Mehta, S. S., & Burks, T. F. (2014). Vision-based control of robotic manipulator for citrus harvesting. *Computers and Electronics in Agriculture, 102*, 146–158.
Mehta, S. S., & Burks, T. F. (2016). Multi-camera fruit localization in robotic harvesting. *IFAC-PapersOnLine, 49*(16), 90–95.
Mehta, S. S., Ton, C., Asundi, S., & Burks, T. F. (2017). Multiple camera fruit localization using a particle filter. *Computers and Electronics in Agriculture, 142*, 139–154.
Quispe, A. H., Milville, B., Gutiérrez, M. A., Erdogan, C., Stilman, M., Christensen, H., et al. (2015, May). Exploiting symmetries and extrusions for grasping household objects. *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3702–3708). IEEE.
Reis, M. J., Morais, R., Peres, E., Pereira, C., Contente, O., Soares, S., … Cruz, J. B. (2012). Automatic detection of bunches of grapes in natural environment from color images. *Journal of Applied Logic, 10*(4), 285–290.
Schiebener, D., Schmidt, A., Vahrenkamp, N., & Asfour, T. (2016, October). Heuristic 3d object shape completion based on symmetry and scene context. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 74–81). IEEE.
Schöler, F., & Steinhage, V. (2015). Automated 3D reconstruction of grape cluster architecture from sensor data for efficient phenotyping. *Computers and Electronics in Agriculture, 114*, 163–177.
Silwal, A., Davidson, J. R., Karkee, M., Mo, C., Zhang, Q., & Lewis, K. (2017). Design, integration, and field evaluation of a robotic apple harvester. *Journal of Field Robotics, 34*(6), 1140–1159.
Thrun, S., & Wegbreit, B. (2005, October). Shape from symmetry. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2* pp. 1824–1831). IEEE.
Vitzrabin, E., & Edan, Y. (2016). Changing task objectives for improved sweet pepper detection for robotic harvesting. *IEEE Robotics and Automation Letters, 1*(1), 578–584.
Wang, X., Ang, M. H., Jr., & Lee, G. H. (2020). Cascaded Refinement network for point cloud completion. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 790–799).
Wang, Z., Walsh, K. B., & Verma, B. (2017). On-tree mango fruit size estimation using RGB-D images. *Sensors, 17*(12), 2738.
Xiong, Y., Ge, Y., Grimstad, L., & From, P. J. (2020). An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation. *Journal of Field Robotics, 37*(2), 202–224.
Xiong, Y., Ge, Y., Liang, Y., & Blackmore, S. (2017). Development of a prototype robot and fast path-planning algorithm for static laser weeding. *Computers and Electronics in Agriculture, 142*, 494–503.
Xiong, Y., Peng, C., Grimstad, L., From, P. J., & Isler, V. (2019). Development and field evaluation of a strawberry harvesting robot with a cable-driven gripper. *Computers and Electronics in Agriculture, 157*, 392–402.
Yang, Q., Chang, C., Bao, G., Fan, J., & Xun, Y. (2018). Recognition and localization system of the robot for harvesting Hangzhou White Chrysanthemums. *International Journal of Agricultural and Biological Engineering, 11*(1), 88–95.
Yin, H., Chai, Y., Yang, S. X., & Mittal, G. S. (2009, December). Ripe tomato recognition and localization for a tomato harvesting robotic system. *2009 International Conference of Soft Computing and Pattern Recognition* (pp. 557–562). IEEE.
Yuan, W., Khot, T., Held, D., Mertz, C., & Hebert, M. (2018, September). Pcn: Point completion network. In *2018 International Conference on 3D Vision (3DV)* (pp. 728–737). IEEE.