Automated flower counting from partial detections: Multiple hypothesis tracking with a connected-flower plant model

Computers and Electronics in Agriculture 188 (2021) 106346

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Automated flower counting from partial detections: Multiple hypothesis tracking with a connected-flower plant model

W. Houtman [a,*], A. Siagkris-Lekkos [a,1], D.J.M. Bos [a], B.J.P. van den Heuvel [b], M. den Boer [b], J. Elfring [c,a], M.J.G. van de Molengraft [a]

[a] Department of Mechanical Engineering, Eindhoven University of Technology, PO Box 513, 5600MB Eindhoven, the Netherlands
[b] Aris B.V., Esp 300, 5633 AE Eindhoven, the Netherlands
[c] Product Unit Autonomous Driving, TomTom, 1011 AC Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Multiple Target Tracking
Multiple Hypothesis Tree
Adaptive Motion Model
Flower Counting
Phalaenopsis Plants

ABSTRACT

This paper presents an automated flower counting method based on Multiple Hypothesis Tracking (MHT) with a connected-flower plant model which is based on detections of flowers. Multiple viewpoints of each plant are taken into account as plants are considered in which flowers can occlude each other. To prevent double counting and to solve inconsistencies caused by false flower detections, a model is developed which describes the plant movement with respect to the camera. The uncertainty of the flower detections is considered in this model. To address variations in the velocity of the plant movement, the model realized in this work explicitly takes into account that motions of flowers are correlated since the flowers are connected to each other via the stem of the plant. This is in contrast to the traditional MHT approach where the movement of each object is typically modeled and estimated separately. In our approach, based on the set of detected flowers, the uncertainty of the plant movement is reduced. As a result, the movement of modeled but not always observed flowers is still properly tracked. To demonstrate the validity of the approach, the proposed counting method is tested on a dataset obtained in a real greenhouse containing multiple viewpoints of 71 Phalaenopsis plants and compared to existing methods. The methods considered include a single viewpoint approach, a heuristic state of the practice approach and an MHT approach with both an independent and connected object description. Within a margin of 1 flower, these methods respectively counted the number of flowers in 44%, 58%, 70% and 92% of the plants correctly. As a result, this work validates the superiority of the MHT approach with a connected-flower plant model.

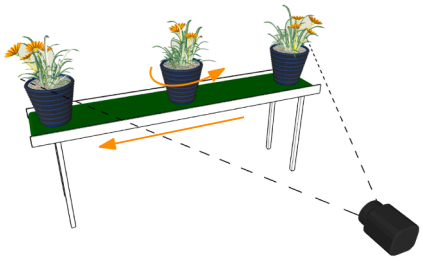
1. Introduction

Automation in the field of agriculture, horticulture and the food industry is increasing rapidly and gradually substituting human labor with machinery. Traditional manual tasks such as harvesting, inspection, sorting and packaging are becoming more and more automated with the ultimate goal being to eliminate the need for human assistance. Among the advantages of automation are effort, time and production cost reduction, availability as well as improved quality and reliability (Mahmud et al., 2020).

An example of such a task is the classification of plants in general and Phalaenopsis plants specifically. The grading of these plants, of which an example is shown in Fig. 1c, is based on a number of features like the height, the number of stems, the number of flowers and buds as well as the color and pattern of the flowers. The quantification of the number of flowers of the Phalaenopsis plant based on computer vision methods is challenging as flowers might be (partially) unobservable due to occlusions. Furthermore, detection inaccuracies of flowers might lead to false detections, missed detections, flower fragmentations and flower merges. In the case of flower fragmentations multiple detections which originate from a single flower are obtained, while for flower merges two or more overlapping flowers are detected as one. Such detection inaccuracies can lead to incorrect estimates of the number of flowers. Therefore, this work focuses on the correct quantification of the number of flowers based on a set of images which are recorded from multiple viewpoints of the same plant. This work is applied to the system which can be seen in

* Corresponding author.
E-mail address: w.houtman@tue.nl (W. Houtman).
[1] Present address: Product Development - Automated Guided Vehicles, Lely, 3147 PB Maassluis, The Netherlands

https://doi.org/10.1016/j.compag.2021.106346
Received 24 April 2021; Received in revised form 12 July 2021; Accepted 18 July 2021



(a) Sketch of the plant rotation and translation along a conveyor belt in front of a static camera.

(b) Conveyor belt and camera box (left) as installed in a greenhouse.

(c) Phalaenopsis plant as recorded by the acquisition setup.

Fig. 1. Conveyor belt system with the acquisition setup to grade Phalaenopsis plants. Within the range of the camera, images are taken as the plant translates and rotates along the conveyor belt.

Fig. 1. Multiple viewpoints of a plant are obtained by moving this plant by means of a conveyor belt in front of a static camera. In order to solve the detection inaccuracies for this specific system, within the current state of the practice the number of flowers is estimated using a heuristic method. In this heuristic method, the images are ranked from high to low based on the number of flower detections. The number of flowers on the image with the third highest number of flowers is considered as the estimated number of flowers. This leads however to inaccurate results. Improvements are sought in a better description of the relations between the consecutive images.

By considering multiple viewpoints of a plant, the same flower is being detected in different images. To derive a correct estimation of the number of flowers, it is desired to know which detections originate from the same flower. Determining which detections come from the same source is within the field of robotics and surveillance known as the data-association problem. This problem can be solved by Multiple Target Tracking (MTT) algorithms (Robin and Lacroix, 2016). Within this work, an MTT algorithm will be selected. These algorithms typically consider objects which are supposed to move independently, such as pedestrians in an automotive context (Luo et al., 2021). In contrast, the flowers of a plant are connected by the stem of the plant and the pot it is placed in, hence motions of flowers are correlated due to motions of the pot and connections to their stem(s). Therefore, within this work a model is created which describes the connected flower-positions within the plant as well as the translation and rotation of the plant over the conveyor belt. We refer to this model as the Connected-Flower Plant Model (CFPM). Due to these interconnections, the entire plant state can be updated despite a subset of the flowers being observed at each image: if for example several flowers are occluded, the rotation of the pot can still be properly estimated based on flowers which are detected. To demonstrate the benefits of this CFPM, we compare this approach to the typical MTT approach where each flower would be modeled independently. This latter model is referred to as the Unconnected-Flower Plant Model (UFPM) and allows for more or less independent flower movements. The flower estimation algorithm is developed by requiring it to (1) solve the occlusions by considering multiple viewpoints, (2) solve the inconsistencies caused by false and missed detections, (3) take into account the detection uncertainty caused by deviations of the detections from the actual flower position and the lack of depth perception, (4) take into account the rotational uncertainty caused by slip and wear of the conveyor belt system and (5) estimate the number of flowers with a maximum allowed deviation of ± 1 flower for each plant. The flower estimation algorithm is validated on a dataset consisting of 71 plants and compared to both a strategy which determines the number of flowers based on a single image and a strategy which estimates the position of each flower in a plant independently. Hereby, the benefit of the CFPM applied on an MHT to consider multiple viewpoints is demonstrated.

The remainder of this paper is structured as follows: the next section gives a literature overview and lists the contributions of this work. Then, Section 3 elaborates on the conventions applied and the methods tested in this work. An overview of the independent flower model as well as the combined plant approach is given. Further, the adaptations of the MTT algorithms are elaborated. Next, in Section 4 these methods are experimentally validated and compared. Section 5 concludes this paper.

2. Related Work

This section discusses the work related to the quantification of objects based on (a set of) images with a focus on approaches considered within agriculture. Next, the multiple target tracking problem will be considered in the context of tracking flowers within a plant with multiple occlusions. The final paragraph summarizes the contributions of this work.

Within the agricultural and horticultural field, quite some work is conducted to automated flower and fruit identification, quality estimation and quantification. Examples are seen in the identification of plant species (Tripathi and Maktedar, 2020), localization of fruits for automated harvesting (Wei et al., 2018; Tang et al., 2020), recognition of fruit diseases (Dubey and Jalal, 2015) and yield prediction (Wang et al., 2013; Stein et al., 2016). As described by Stein et al. (2016), progress over the last years from the machine vision community with convolutional neural networks has led to highly accurate fruit detection in color imagery. Some recent overviews of these developments can be found in Koirala et al. (2019), Tripathi and Maktedar (2020), Wäldchen and Mäder (2018), Naranjo-Torres et al. (2020). For counting purposes, it is typically seen that these techniques are applied to determine the number of fruits or flowers based on a single image. Examples of this approach can be found in Bairwa and Agrawal (2014), Sethy et al. (2019), Syal et al. (2014), Hamidinekoo et al. (2020). These methods assume all fruits or flowers to be visible on a single image as occlusions can not be taken into account. Though in an object detection context the handling of occlusions is addressed, these methods are not robust for heavy or full occlusions (Wang et al., 2017; Koporec and Pers, 2019; Yang et al., 2021). As a result, for the purpose of counting flowers in the Phalaenopsis plant, quantifying the number of flowers based on a single view is not possible as significant parts of the flowers might be (fully) occluded such as can be observed in Fig. 1c. Furthermore, false detections, both positive and negative, can not be resolved based on a single image. Therefore, within this work, multiple images need to be considered with multiple viewpoints of the plant which requires to determine which

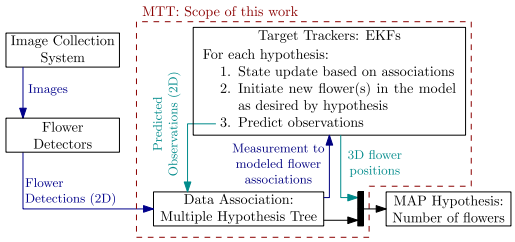


Fig. 2. Overview of the proposed MTT algorithm.

detections in the dataset originate from the same flower.

To derive which detections in the dataset originate from the same flower, the Multiple Target Tracking problem for an *a priori* unknown number of targets needs to be resolved. This problem is widely studied in the field or robotics and surveillance, leading to several approaches which each come with its possibilities and limitations (De Laet, 2010; Robin and Lacroix, 2016). The MTT-problem typically combines data association and model-based object tracking (Elfring et al., 2013), which both will be discussed.

The first component of the MTT considers the data-association problem. Data-association algorithms typically try to consider the possible explanations of uncertain object measurements by determining which measurement belongs to which of the objects modeled and their corresponding tracks. Typically, possible measurement-to-object associations are taken into account in a hypothesis. Variations with an increasing (computational) complexity are (1) the Global Nearest Neighbor (GNN) method, (2) the Joint Probabilistic Data Association Filter (JPDAF) and (3) the MHT approach (Elfring et al., 2013). The GNN-approach is within an agricultural context successfully applied by means of the Hungarian algorithm to track mangoes on an orchard (Stein et al., 2016). Within the agricultural context, similar ideas are proposed by Harmsen and Koenderink (2009); Song et al. (2014); Moonrinta et al. (2010). The GNN-method assumes that a measurement can originate from one target or from clutter and that a target can produce zero or one measurement at a time. Associations are based on the current measurement and even in the case of ambiguities no alternatives are maintained. This leads to irreversible decisions as assignments cannot be influenced by data collected in later measurements (de Waard, 2008). Consequently, this method is not robust against target occlusions (De Laet, 2010). Therefore, the GNN-approach is not suitable for the quantification of the flowers of a Phalaenopsis plant as considered in this work. In contrast, the MHT-approach considers all possible explanations or hypotheses from measurements to objects in parallel. As plausible alternatives are maintained, based on new evidence hypotheses could switch since a more probable explanation of the object tracks is available. The JPDAF results can be obtained from an MHT by representing the world by a probabilistically weighted average of a subset of hypotheses and again prune all other hypotheses (Elfring et al., 2013). Like the GNN approach, irreversible decisions are made. For the flower counting problem it is however not required to, at each point in time, to have a single estimate. The most probable explanation is sought after having received all measurements and switching between hypotheses is not a problem. As the MHT-approach is capable in considering the plausible explanations in parallel, this approach will be selected for this work. A typical problem of MHT is that this approach is both exponential in time and memory such that in practice only a limited number of hypotheses can be maintained. In this work, soft real-time constraints of $3-4$ [s] are considered, which indicates the typical time in which a plant moves from the image collection system to the sorting system. Within this time horizon, it is assumed that a sufficient set of alternatives can be evaluated and hence, the time and memory problems are not considered as an issue. Concretely, the MHT as described by Elfring et al. (2013) and available at Elfring (2021) is adopted and modified within this context.

Whereas the first step of the MTT hypothesized about the possible measurement-to-object associations, the second component of the MTT-problem considers the model-based object tracking. In the latter component, each measurement is used to refine the object's state estimate. For MTT-problems, typically detections are associated with objects that are considered to move independently. Within an agricultural context, (Harmsen and Koenderink, 2009) applied a particle filter to estimate the ellipsoidal projection on an image for each flower of a plant rotating in front of a camera. A set of particles that are hypotheses for the flower position in the image plane represent the posterior probability distribution. In combination with the MHT for solving the data-association problem, this leads to a combinatorial explosion of hypotheses. This makes the particle filter not suitable for our approach. Furthermore, this reference considers each flower independently given a measured rotation of the plant. In contrast, within the proposed work unknown but bounded variations of the rotational velocity are considered as the rotational velocity is not directly measured. However, the rotational velocity is observable as the detections of the set of flowers are correlated since the plant translates and rotates as a single entity over the conveyor belt. To demonstrate the benefits of taking this correlated movement into account, an MHT-approach which, similarly to Harmsen and Koenderink (2009), models each flower as a separate object is compared to an MHT-approach which considers the correlated movement. For the latter alternative, a state representation is applied which represents the positions of all flowers with respect to the pot center as well as the translational and rotational velocity of the pot center. Note that depending on the number of flowers considered within a hypothesis, the state size varies. For this approach, the MHT-tracker is revised such that associations can be made based on the flowers of the plant in order to update the estimate of the plant movement given these associations.

In conclusion, the contributions of this work are:

- A method which estimates the number of flowers in a single plant by means of an MHT-tracker. In comparison with the single image approach and a heuristic state of the practice method, it will be demonstrated that by consistently considering multiple views, occlusions and false detections are largely resolved.
- A revision of the MHT-approach where partial detections contribute to a single object model by considering the relations between the object model and the partial detections. Within the context of this work, this means that the state contains the pot movement and all flower positions with respect to the pot and the measurements contain the flower positions. Here, the stem is assumed to be static. As a result, correlations between estimated flower positions and the plant movement can be exploited such that a consistent and more accurate plant movement is obtained.
- The validation of and a comparison between the connected and unconnected-flower plant model, or generally speaking, a connected and independent object description by means of a dataset obtained in a greenhouse.

3. Methods

This section elaborates on the proposed algorithm. An overview is shown in Fig. 2. As indicated in this figure, a system to collect a set of images with multiple viewpoints of a plant and a method to detect the positions of the individual flowers on the image is considered as a prerequisite. For each image, these flower detections are fed to the MHT to (1) associate the detected flowers to flowers previously modeled, (2) initiate a new flower in the model or (3) mark the detection as clutter. These alternatives are considered as the association hypotheses in the data-association step of the figure. The next step is to maintain each hypothesis, which is the task of the "Target Tracker". Given the associations corresponding to each hypothesis, the plant state, which represents the (rotational) movement of the plant and the corresponding

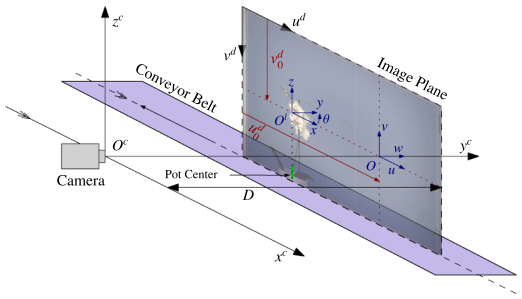


Fig. 3. Graphical representation and conventions of the image collection system.

flower positions, is estimated and maintained within a single state vector. Based on the expected plant motion, the flower positions for the subsequent frame are predicted. By comparing the predicted detections to the actual detections, a new set of hypotheses is formed and validated by determining the probability of a hypothesis being correct. Finally, when all images are processed, the state of the Maximum A Posteriori (MAP) hypothesis indicates the number of flowers present in a plant. After discussing the conventions, the remainder of this section elaborates on each of the MTT components. Considering the target tracker, the object description where each flower is designed as an independent object as well as the description where all flowers are considered as a component of the object description with both a correlated movement and a correlation between the object positions will be elaborated. Regarding the data association algorithm, the consequences of the object description will be discussed. For clarity, an overview of the nomenclature as applied in this work is given in Table 4 of Appendix A.

3.1. Conventions and Flower projections onto the Image Plane

In this paragraph, the conventions as applied in this work are discussed. Based on this discussion, the projections on the image plane are derived in the second part of this section. These projections are used in the observation models of the target tracker.

A graphical representation of the image collection system is provided in Fig. 3. In this figure, a sketch of the camera and the conveyor belt is provided. The image plane is projected at the center of the conveyor belt, along its motion direction. The flower detections, which are represented by a circular bounding box of which its center indicates the center of each flower, are expressed in the image-frame (u^d, v^d). Further, the u^d-coordinate of the detected pot center is measured. The image-center $(u_0^d,\ v_0^d)$, expressed with respect to the image-frame, is determined by calibration. Here, the (u, w, v)-frame is located. The camera is located at the origin of the (x^c, y^c, z^c)-frame, while the image plane is found along the center of the conveyor belt at a distance D in the y^c-direction of the (x^c, y^c, z^c)-frame. As a result, the origin of the (u, w, v)-frame expressed with respect to O^c equals $(0, D, 0)$. The flower-coordinates (x, y, z) are represented in a local frame with its origin O^l above the pot center. Hence, given the pot center location u expressed in the (u, w, v)-frame, the translation from the (u, w, v)-frame to the flower frame (x, y, z) equals $(u, 0, 0)$. The rotation around the z-axis is represented by θ, with the rotational velocity ω equal to its time-derivative, hence $\omega = \dot{\theta}$.

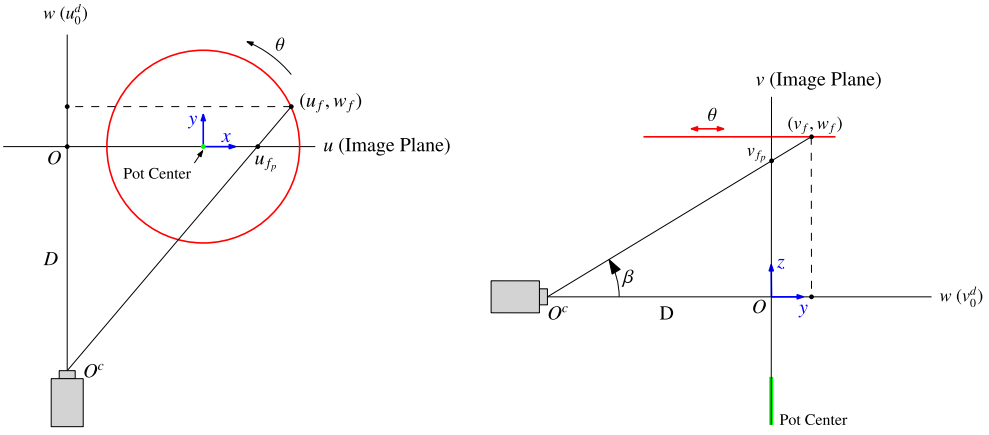
Given the conventions, the projections of the flowers on the image-plane can be derived based on the principles of parallel projection (Spong et al., 2005). Hereby it is assumed that lens-distortion is negligible or corrected. Fig. 4 shows a top view and a side view of a flower f having position (u_f, w_f, v_f) expressed in the (u, w, v)-frame and its projection f_p on the image having coordinates $(u_{f_p}, 0, v_{f_p})$. Given that $\triangle O^c$, (u_f, w_f), $(0, w_f)$ and $\triangle O^c$, $(u_{f_p}, 0)$, O of Fig. 4a as well as the $\triangle O^c$, (v_f, w_f), $(0, w_f)$ and $\triangle O^c$, $(v_{f_p}, 0)$, O of Fig. 4b are similar, the following set of relations is derived:

$$\left.\begin{aligned} \frac{u_f}{u_{f_p}} &= \frac{w_f + D}{D} \\ \frac{v_f}{v_{f_p}} &= \frac{w_f + D}{D} \end{aligned}\right\} \leftrightarrow \left.\begin{aligned} u_{f_p} &= \frac{u_f}{1 + w_f/D} \\ v_{f_p} &= \frac{v_f}{1 + w_f/D} \end{aligned}\right\}. \tag{1}$$

3.2. Target Tracking: Unconnected versus Connected Flower Model

Next, the model-based target tracking component as shown in Fig. 2 is discussed. The task of this component is to update the object state estimation based on the associations as determined by the hypotheses tree. Both the UFPM as well as the CFPM are presented. Further, for the hypotheses where new flowers are considered, in both cases an initialization of the three dimensional flower position is required based on a two dimensional detection on the image plane. The principles applied for this initialization step will be discussed.

Under the assumption of additive zero mean Gaussian measurement noise in flower detections on the image plane and a (non) linear process model associated with Gaussian process model noise in the (rotational) movement of the plant, for both the independent flower representation as well as the combined plant representation, an approach based on an Extended Kalman Filter (EKF) is proposed. For clarity and notation, the process and observation equations are provided here. More details on (E)KFs can for example be found at Thrun et al. (2005); Welch and Bishop (2006). When considering a process model with no control input,



(a) Top view of the flower movement. (b) Side view of the flower movement.

Fig. 4. Top and side view of the flower, while the plant rotates on the conveyor belt.

the process model and the measurement model are described by the following equations:

$$\mathbf{x}_k = g(\mathbf{x}_{k-1}) + \mathbf{q}_k \tag{2}$$

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{s}_k. \tag{3}$$

Here, the estimate $\hat{\mathbf{x}}$ of the state vector $\mathbf{x} \in \mathbb{R}^n$ at discrete time step k with dimensions n is governed by a nonlinear process model g and the zero mean Gaussian process noise vector $\mathbf{q}$. A measurement $\mathbf{z} \in \mathbb{R}^m$ of dimension m is described by a nonlinear observation model h and the zero mean Gaussian measurement noise vector $\mathbf{s}$. The zero mean covariance matrices of both noise variables are denoted as Q and S, respectively. At every time step, a state estimate can be obtained using a two step method, namely a prediction step and an update step. For details about these steps, the reader is referred to Thrun et al. (2005); Welch and Bishop (2006). Further, the Jacobians of the non-linear functions g and h evaluated at a given state estimate are respectively represented by G_k and H_k.

3.2.1. Connected-Flower Plant Model

For the CFPM, the movement of the plant along the conveyor belt as well as the entire set of flowers is represented in the state vector. The plant movement is described by the translational and rotational velocity of the pot center and each flower is described by its position with respect to this pot center. For the local frame (x, y, z), the first measurement is chosen as reference configuration. Since the flowers present within a plant are considered to have a fixed configuration, this frame is considered static. As a result, when omitting time step k in the notation, the plant state $\mathbf{x}_P$ is represented as

$$\mathbf{x}_P = \begin{bmatrix} \omega & u_0 & \dot{u} & \mathbf{x}_{f_1}^T \cdots \mathbf{x}_{f_N}^T \end{bmatrix}^T. \tag{4}$$

Here, the first three states represent the plant movement. These consist of the (average) rotational velocity ω of the pot center since the initial detection, the initial position of the pot center at the first image u_0 and the (average) translational velocity of the pot center $\dot{u}$. Also, the estimation of the translational pot center velocity is considered since the initial detection. The latter two variables are expressed with respect to the center frame (u, w, v). The remaining elements of the state vector in (4) represent the estimated start position $\mathbf{x}_{f_i}$ of flowers $f_i, i = 1, \ldots, N$ if the (rotational) velocity of the pot would have been constant. Here, N indicates the number of flowers considered in the plant state description of (4). Note that different hypotheses may have different values for N and, in general, will have different flower position estimates $\mathbf{x}_{f_i}$. The flower states are described by the Cartesian flower position with respect to the pot center, hence the estimated start position (x_0, y_0, z_0) is expressed in the local plant frame. This gives

$$\mathbf{x}_{f_i} = \begin{bmatrix} x_{i,0} & y_{i,0} & z_{i,0} \end{bmatrix}^T. \tag{5}$$

As the rotational and translational velocity of the pot center are assumed to be constant and the relative position of the flowers within the plant is considered static, the state transition represents a constant process model g_P, hence

$$\mathbf{x}_{P,k|k-1} = g_P\left(\mathbf{x}_{P,k-1|k-1}\right) = \mathbf{x}_{P,k-1|k-1}. \tag{6}$$

The predicted state at time step time k given the state estimate up to the previous time step $k-1$ is denoted by subscript as $k|k-1$. The Jacobian G_P of the estimated state equals the identity matrix I, giving

$$G_P = I \in \mathbb{R}^{3(N+1)\times 3(N+1)}. \tag{7}$$

The expected measurement vector $\mathbf{h}$ at time-step k considers the expected pot center position as well as the expected measurement of each flower contained in $\mathbf{x}_P$ given the pot center movement. As a result, when considering M flowers being associated with measurements received at time k, the expected measurement vector of the plant equals

$$\mathbf{h}_P^T(\mathbf{x}_{P,k|k-1}) = \begin{bmatrix} \mathbf{h}_{f_1}^T(\mathbf{x}_{P,k|k-1}) & \cdots & \mathbf{h}_{f_M}^T(\mathbf{x}_{P,k|k-1}) \end{bmatrix}^T. \tag{8}$$

Note that in the prediction step there are no associations to the measurements yet, hence $M = N$ in this step. In the equation, the expected flower measurement h_{fi} of flower fi is represented by

$$\mathbf{h}_{f_i}\left(\hat{\mathbf{x}}_{P,k|k-1}\right) = \begin{bmatrix} \hat{u}_{k|k-1} \\ \hat{u}_{pf_i,k|k-1} \\ \hat{v}_{pf_i,k|k-1} \end{bmatrix} = \begin{bmatrix} \hat{u}_{0,k|k-1} + \hat{\dot{u}}_{k|k-1} \cdot dt \\ \hat{u}_{pf_i,k|k-1} \\ \hat{v}_{pf_i,k|k-1} \end{bmatrix}. \tag{9}$$

The first element of this vector indicates the expected position of the pot center and is based on the position of the pot center at the initial frame, its estimated translational velocity and the time dt between the measurement at time step k and the initial frame. The second and third elements of this vector consider the estimated projections of the flowers $(\hat{u}_{pf_i}, \hat{v}_{pf_i})$. These are determined according to (1) and the expected flower position in the (u, w, v)-frame. This expected flower position is estimated based on the translation and rotation of the plant. It is determined as

$$\begin{bmatrix} \hat{u}_{f_i,k|k-1} \\ \hat{w}_{f_i,k|k-1} \\ \hat{v}_{f_i,k|k-1} \end{bmatrix} = \begin{bmatrix} \hat{u}_{0,k|k-1} + \hat{\dot{u}}_{k|k-1} \cdot dt \\ 0 \\ 0 \end{bmatrix} + R\left(\hat{\theta}_{k|k-1}\right) \mathbf{x}_{f_i,k|k-1}. \tag{10}$$

Based on the image update frequency f, the expected rotation $\hat{\theta}_{k|k-1}$ is determined as

$$\hat{\theta}_{k|k-1} = \hat{\omega}_{k|k-1} \cdot dt = \hat{\omega}_{k|k-1} \cdot k \Big/ f, \tag{11}$$

while $R(\theta)$ represents the rotation matrix around the pot center or z-axis. Note here that the EKF is initialized at $k = 0$. The rotation matrix is computed as

$$R\left(\hat{\theta}_{k|k-1}\right) = \begin{bmatrix} \cos\left(\hat{\theta}_{k|k-1}\right) & -\sin\left(\hat{\theta}_{k|k-1}\right) & 0 \\ \sin\left(\hat{\theta}_{k|k-1}\right) & \cos\left(\hat{\theta}_{k|k-1}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

The process covariance matrix quantifies the uncertainty in the process model of $\mathbf{x}_P$. Within this context, causes for uncertainty are for example by slip in the pot center movement and vibrations of the flowers during their movement on the conveyor belt. The covariance depends on the number of flowers and is considered as

$$Q_{P,k} = \begin{bmatrix} Q_{PC,k} & 0^{3\times 3N} \\ 0^{3N\times 3} & \mathrm{diag}\left(\begin{bmatrix} Q_{f_1,k} \ldots Q_{f_N,k} \end{bmatrix}\right) \end{bmatrix}. \tag{13}$$

By using a block-diagonal structure, we assume that the uncertainty in the process model between the pot center and all flowers is uncorrelated. The covariance of the process model noise of the pot center movement is indicated by $Q_{PC} \in \mathbb{R}^{3\times 3}$ and the flower process covariance of flower i is represented by $Q_{f_i} \in \mathbb{R}^{3\times 3}$. In a similar way, the measurement error covariance matrix of the measurement vector of the plant (8) is modeled. It is determined as

$$S_{P,k}^{3M\times 3M} = \mathrm{diag}\left(\begin{bmatrix} S_{f_1,k} & \ldots & S_{f_M,k} \end{bmatrix}\right). \tag{14}$$

3.2.2. Unconnected-Flower Plant Model

For the UFPM, the reasoning is similar as for the CFPM, but a maximum of one flower is considered in each state, hence $N = 1$. Consequently, each hypothesis has multiple EKFs, namely one for each flower, whereas in CFPM one EKF per hypothesis is sufficient. As a result, for the unconnected-flower plant representation $\mathbf{x}_f$ the plant state description of (4) reduces to

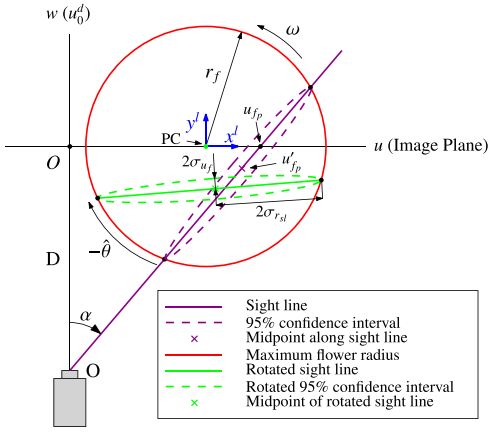


Fig. 5. Top view of the estimation of the initial flower position and its covariance based on the maximum flower radius.

$$\mathbf{x}_f = \begin{bmatrix} \omega & u_0 & \dot{u}\,\mathbf{x}_{f_1}^T \end{bmatrix}^T. \tag{15}$$

Similar to the CFPM, the rotational velocity, the initial position of the pot center and the translational velocity of the pot center are described. However, as indicated by the last element of the vector, for the UFPM a single flower is considered. As a result, the initial and the (rotational) velocity of the pot are estimated for each flower separately thereby allowing independent flower movements. For the UFPM, the state transition model g_f, the observation model h_f, the process covariance matrix $Q_{f,k}$ and observation covariance matrix $R_{f,k}$ are evaluated similarly to the plant representation, thereby considering $N = 1$. As a result, the flower transition model describes a constant state, giving

$$\mathbf{x}_{f,k|k-1} = g_f\left(\mathbf{x}_{p,k-1|k-1}\right) = \mathbf{x}_{f,k-1|k-1}. \tag{16}$$

Its Jacobian G_f equals the identity matrix, hence

$$G_f = I \in \mathbb{R}^{6\times 6}. \tag{17}$$

The predicted flower observations are determined as

$$\mathbf{h}_f\left(\mathbf{x}_{f,k|k-1}\right) = \mathbf{h}_{f_1}\left(\mathbf{x}_{f,k|k-1}\right), \tag{18}$$

in which $h_{f_1}(\mathbf{x}_{f,k|k-1})$ is computed according to (9)–(12). Similar to (13) and (14), the process covariance matrix and the observation covariance matrix are for the UFPM respectively derived as

$$Q_{f,k} = \begin{bmatrix} Q_{PC,k} & 0^{3\times 3} \\ 0^{3\times 3} & Q_{f_1,k} \end{bmatrix} \tag{19}$$

and

$$S_{f,k} = S_{f_1,k}. \tag{20}$$

3.2.3. Initialization of Flower State

When initializing a new flower in the state representation, for both the CFPM and UFPM, the full flower state and its corresponding covariance must be initialized based on the measurement on the image plane. In both cases the state initialization is based on this measurement and the maximum radius with respect to the pot center r_f at which a flower is typically found. An overview is depicted in Fig. 5. Here a top view of the situation is shown with respect to the center frame. A projected flower f_p which is detected on the image plane at $(u_{f_p}, 0, v_{f_p})$ is considered. The sight line between the camera is drawn in magenta and the maximum flower radius with respect to the pot center is drawn in red. The initial flower position is determined as the center of the intersections u'_{f_p} between the sight line and the maximum flower radius r_f at which the flowers are found. In the figure, this initial flower position is shown with a magenta cross. As the flower positions are represented with respect to the pot center at the initial frame, the initial position vector $\mathbf{x}_{f_i,0}$ of flower f_i is required with respect to the moment of state initialization. Therefore, u'_{f_p} is rotated according to the estimated rotation as

$$\mathbf{x}_{f_i,0} = R\left(-\hat{\theta}_{k|k}\right)u'_{f_p}. \tag{21}$$

This derivation leads to the position represented with a green cross in Fig. 5. Note that for the individual flower representation the initial position is estimated at the moment of detection, hence there is no rotation considered. As a result, for this situation $R(-\hat{\theta}_{k|k}) = I \in \mathbb{R}^{3\times 3}$. Further, θ is considered at $k|k$. This means that based on the associated measurements, first the update step is applied before the state is extended as this gives a more accurate estimate.

The covariance Λ with respect to the initial center point u'_{f_p} is drawn with a magenta dotted line and determined along the sight line as

$$\Lambda = \mathrm{diag}\left(\begin{bmatrix} \sigma_{u_f}^2 & \sigma_{r_{sl}}^2 & \sigma_{v_f}^2 \end{bmatrix}\right). \tag{22}$$

In this equation, the standard deviation with respect to the center point in the direction of the sight line is represented by $\sigma_{r_{sl}}$. This distance equals half the difference between the intersection points. σ_{u_f} and σ_{v_f} represent the standard deviation of the detection uncertainty in the u and v-direction of the image plane respectively. Given that $D \gg r_f$ and $D \gg \left|u^c_{f_p}\right|$, these variables are assumed to be mutually perpendicular to the sight line. To represent the initial flower covariance $P_{f_i,0}$ in the local frame, the rotation between the sight line and the local frame and the flower rotation with respect to the initial position is considered as

$$P_{f_i,0} = R\left(-\hat{\theta}_{k|k}\right)\underbrace{R_0(\alpha,\beta)\Lambda R_0^T(\alpha,\beta)}_{\text{Covariance along sight line}}R^T\left(-\hat{\theta}_{k|k}\right). \tag{23}$$

Here, the initial rotation matrix $R_0(\alpha,\beta)$ indicates the rotation between the sight line and the local frame. It utilizes the angle $\alpha = \mathrm{atan2}\left(u_{f_p}, -D\right)$, visualized in Fig. 5, between the (w,v)-plane and the sight line. In a similar fashion, the angle $\beta = \mathrm{atan2}\left(-D, v_{f_p}\right)$ between the (u,w)-plane and the sight line is applied. This angle is visualized in Fig. 4b. The rotation matrix is computed as

$$R_0\begin{pmatrix}\alpha,\beta\end{pmatrix} = \begin{bmatrix} c(\alpha) & -s(\alpha) & 0 \\ s(\alpha) & c(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -s(\beta) & c(\beta) \\ 0 & c(\beta) & s(\beta) \end{bmatrix}. \tag{24}$$

For the sake of compact notation, $s()$ and $c()$ represent the sine and cosine function respectively. Further, the estimated plant rotation is utilized to obtain the covariance at the moment of the initial detection. As a result, the green dotted line of Fig. 5 is obtained.

3.3. MHT: Adaptations for Single State Representation based on Partial Detections

This section elaborates on the data-association component of the counting algorithm as shown in Fig. 2. The MHT-algorithm as proposed by Elfring et al. (2013, 2021) is adopted. In accordance with the UFPM of the target tracker, each object considered in the MHT is considered independent. As this is in contrast with the CFPM, the adaptations and consequences of the MHT data-association algorithm for the plant representation based on flower detections are discussed.

The main idea of the MHT-framework of Elfring et al. (2013, 2021) is that each measurement can (1) represent an object not contained in the the state vector, (2) originate from a flower which is present in the state vector already or (3) be a false detection. Within the framework applied,

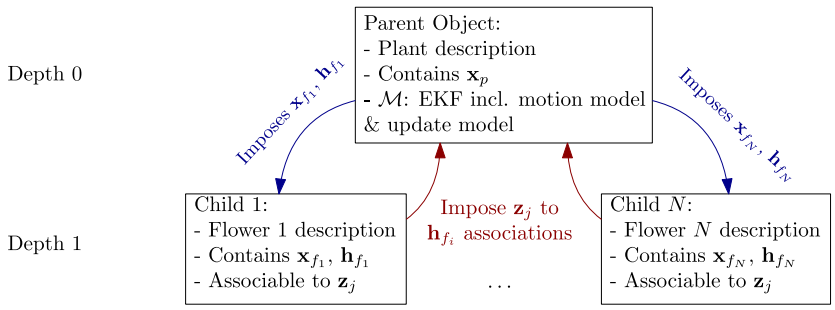


Fig. 6. Visualization of parent–child relationships as applied in the object description for the CFPM of this work. Each hypothesis contains its own instance of this figure.

Table 1
Parameters of the flower counting system as applied in this work.

Parameter	Quantity	Value
D	Distance between camera and image plane	1.808 [m]
f	Image update frequency	13.33 [Hz]
r_f	Maximum radius w.r.t. the pot center at which flowers are found as indicated in Fig. 5	$18{\cdot}10^{-2}$ [m]
(u_0^d, v_0^d)	Image center	$(960, 820)$ [pixels]
σ_{u_f}	Standard deviation in u-direction of flower measurement	0.01 [m]
$u_{max} \times v_{max}$	Image dimensions	1920×1440 [pixels]
$\dot{u}_0$	Expected initial translational velocity of the pot center	-0.35 [m/s]
σ_u	Standard deviation of the pot center measurement and process uncertainty	$5.0{\cdot}10^{-2}$ [m]
$\sigma_{\dot{u}}$	Standard deviation of the pot center velocity along the conveyor belt	$3.2{\cdot}10^{-2}$ [m/s]
σ_{v_f}	Standard deviation in v^c-direction of flower measurement	0.01 [m]
$\sigma_x, \sigma_y, \sigma_z$	Standard deviation of flower process uncertainty in x, y, z-direction	$1{\cdot}10^{-3}$ [m]
σ_{ω_0}	Initial standard deviation of the rotational velocity	$0.025{\cdot}\omega_0$
σ_ω	Standard deviation of the rotational velocity	$0.01{\cdot}\widehat{\omega}$
ω_0	Expected initial rotational velocity of the pot center	2.54 [rad/s]

the general idea is to associate the observations to objects estimated. As such, an object o represents the object as

$$o_a = o_a\left(\iota, \mathbf{z}_k^j, \mathscr{M}_{k,a}\right). \tag{25}$$

In this equation, ι represents a unique identifier, $\mathbf{z}_k^j$ represents a measurement performed at time step k that will be linked to this object and $j = 1, \ldots, M$ is the measurement index of a measurement set containing M measurements. Further, $a = 1, \ldots, N_o$ is the object index of a hypothesis which describes N_o objects. $\mathscr{M}_{k,a}$ contains the behavior model of the object. For the UFPM each flower is modeled independently, hence N_o equals the number of flowers N contained in a hypothesis. $\mathscr{M}$ contains the independent flower $\mathbf{x}_f$ state (15) and the EKF including its process model (16) and observation model (18). As each flower is modeled individually, this description is sufficient for the UFPM. In contrast, for the CFPM this description is insufficient as the motion model describes the plant movement of multiple flowers which are detected separately. To describe the relations between the plant movement and the flower observations as contained in a hypothesis, the object description is extended with a parent–child description according to

$$o_{ap} = o_{ap}\left(\iota, \mathbf{z}_k^j, \mathscr{M}_{k,ap}, o_{ac}\right). \tag{26}$$

In here, $ap = 1, \ldots, N_{po}$ and $ac = 1, \ldots, N_{co}$ respectively represent the index of the parent and child objects contained in a hypothesis. An overview of the relations as applied in this work is shown in Fig. 6. Here it is shown that the parent object describes the plant, namely the CFPM. Since a single plant is contained in a hypothesis, $N_{po} = 1$ for the problem considered in this work. Its behavior model $\mathscr{M}_{k,ap}$ contains the plant state $\mathbf{x}_P$ (4) and both the motion model (7) and observation model (8) of the EKF at time step k. In order to associate the child objects, the flowers in this work, to the flower measurements, the (predicted) flower measurement (9) is imposed by the parent object. In return, the actual measurement to modeled flower object associations of a hypothesis are derived via the child objects. As in this work each flower is described as a child object, $N_{co} = N$.

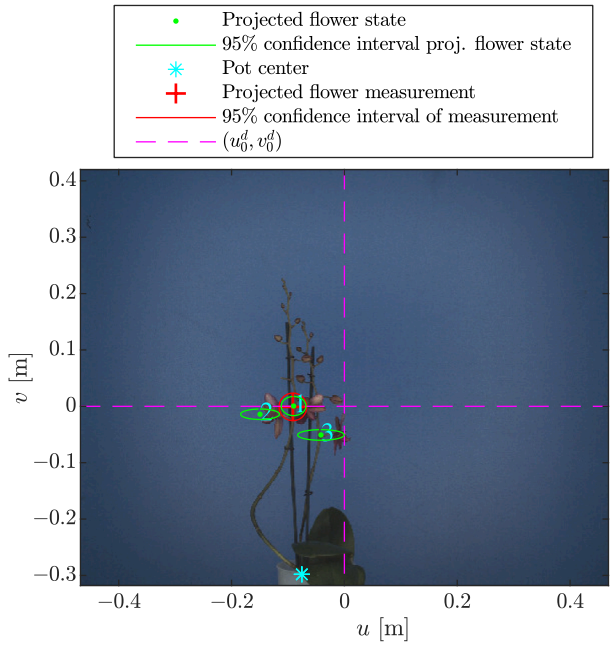
Once the possible associations are determined, the object state can be updated and the probabilities of each hypothesis being correct can be computed. Compared to the single object approach, in accordance with (21), the order of updating objects and initializing new objects becomes relevant for the CFPM as the initialization of new flowers in an existing plant depends on the rotational velocity. To achieve a more accurate estimate of this velocity, for the CFPM, first the update step is applied based on all associations to flowers which are present in the state description $\mathbf{x}_P$. In the second phase, the new objects are initialized based on the estimated rotational velocity. To update the hypotheses probabilities, for both state estimates and object-descriptions Bayes' rule is used as described in Elfring et al. (2013). For a detailed derivation, the reader is referred to Section 5 of this reference. In the context of this work, the Mahalanobis distance between a flower measurement on the one hand and the expected flower projection onto the image plane as derived in (1), (9) and (10) on the other hand is applied as an important criterion to determine the likelihood of an association.

4. Experiments & Discussion

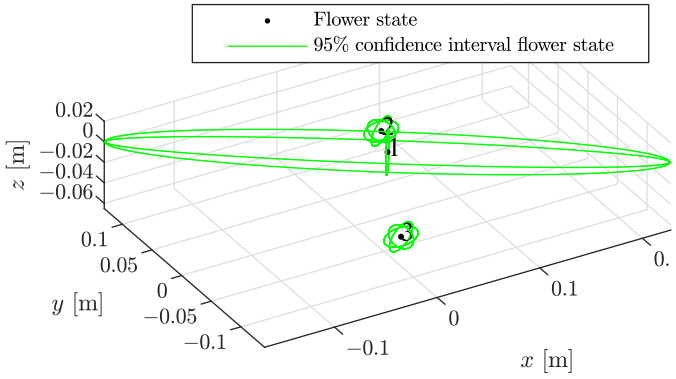
To compare the single image approach with both the CFPM and UFPM present in the MTT for the multiple image approach, this section presents and discusses the results achieved for these methods. First, the system properties and algorithms applied to obtain the flower detections are described. Next, the functionality of the hypothesis tree will be validated for both models, followed by the improvements in the quantification of the number of flowers in each plant. In the final subsection, suggestions for future work are provided.

4.1. System Properties and MHT Settings

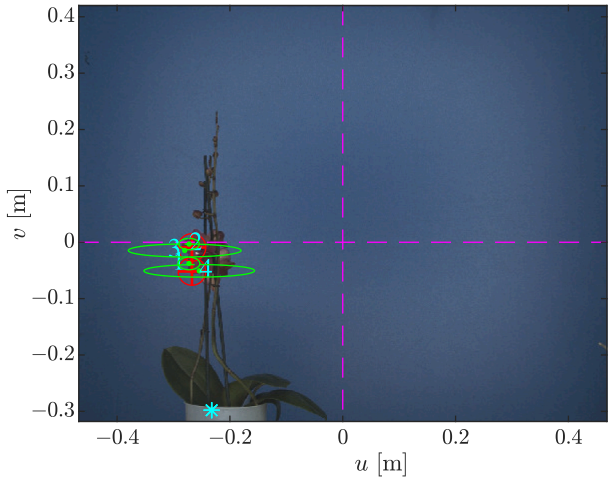
To test the methods developed in this work, a dataset of 71 Phalaenopsis plants was collected. The dataset originates from a real greenhouse and is thus considered as representative for real applications. A Sony XCG-H280CR-camera running at 13.33 [Hz] collected images of the plants and the flower counting algorithm was processed on a HP ZBook Studio G3 with an Intel i7-6700HQ 2.60 GHz × 8 processor running Ubuntu 16.04. Within the camera-view, the plant typically makes a rotation of 210 [degrees] to record a set of 20 images. The



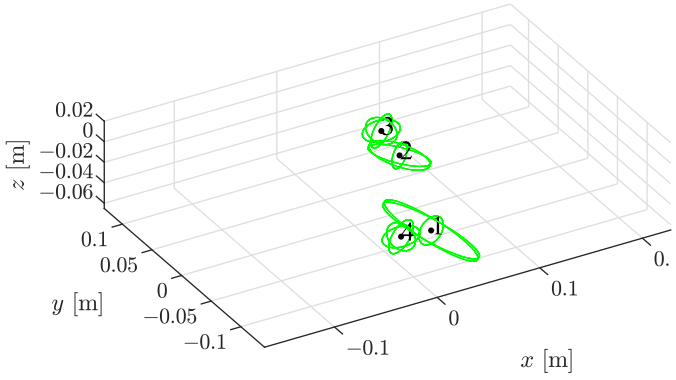
(a) Visualization of the independent model after processing
13 images.



(b) Visualization of the UFPM after processing 13 images. Flower 1 is initialized for the first time.



(c) Projected flower states after processing all 20 images.



(d) Final estimate of the flower states.

Fig. 7. Visualization of the UFPM (right) and the projected flower states on the images (left) after processing the detections. The rotational velocity for each flower is listed in Table 2. The plant consis.ts of 3 flowers.

arrival of a plant is measured with an external sensor which ensures that the full plant is recorded at the initial image. Each plant contained up to 25 flowers. The distribution of the number of flowers over the plants is shown by means of a histogram in Fig. 10i. The flower detections were obtained using local binary patterns and AdaBoost learning as described and validated by Puttemans and Goedemé (2015). Further, for each image the pot center is determined by measuring the edges of the pot based on the color difference between the pot and the background.

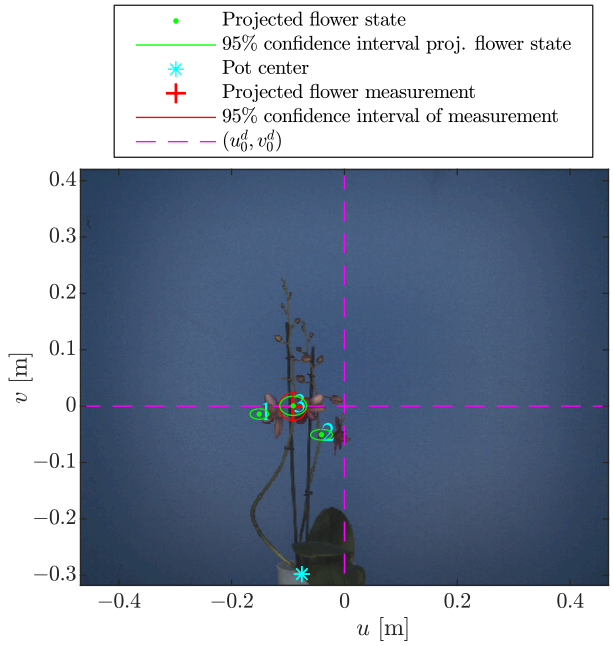
The MTT settings depend on the plant type and setup. An overview of the MTT settings as applied for the Phalaenopsis plants is provided in Table 1. For both the CFPM and the UFPM and of the target tracker, similar settings were chosen. The initial translational $\dot{u}_0$ and rotational velocity ω_0 of the plant are set to the nominal velocities as configured within the image collection system. To account for unmodeled dynamics such as slip of the pot movement and flower vibrations caused by plant accelerations, the process uncertainty of both the CFPM and the UFPM as respectively indicated in (13) and (19) are considered. The process uncertainty $Q_{PC,k}$ of the movement of the pot center at time step k is represented as

Table 2
Estimated rotational velocity [rad/s] of the flower states ±2 standard deviation of Fig. 7.

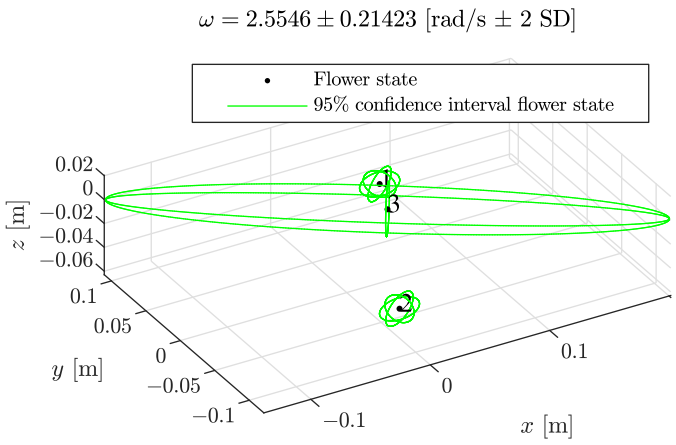
	13 images considered	Full set considered
ω_{f_1}	2.5386 ± 0.1376	2.5369 ± 0.1631
ω_{f_2}	2.5561 ± 0.2162	2.5389 ± 0.1854
ω_{f_3}	2.5408 ± 0.2154	2.5561 ± 0.2494
ω_{f_4}	–	2.5408 ± 0.2487

$$Q_{PC,k} = \mathrm{diag}\left(\left[\sigma_{\omega}^2 \quad \sigma_{u_0}^2 \quad \sigma_{\dot{u}}{}^2\right]\right). \tag{27}$$

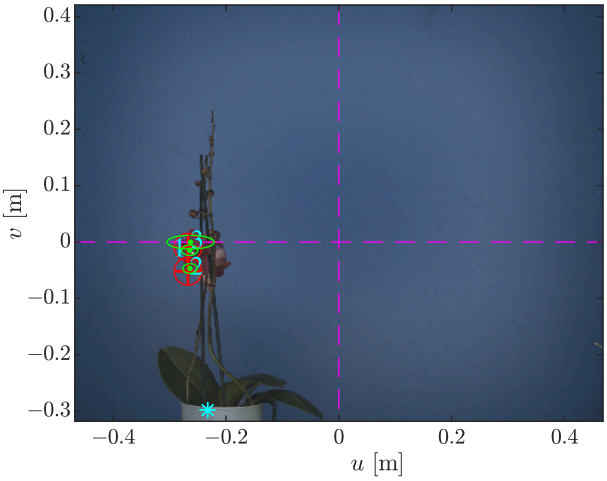
Here, σ indicates the Standard Deviation (SD) of the respective state variable. As indicated by the table, the main source of uncertainty of the plant movement is the rotational velocity caused by slip and wear of the conveyor belt. Based on experience, the deviation σ_{ω_0} is typically bounded within 5% of the rotational velocity and as such this value is chosen as its 95% confidence interval. The main source of uncertainty of the flower positions with respect to the pot center is the occurrence of flower vibrations due to rotational plant acceleration. In between consecutive images, the standard deviation of the flower vibrations is



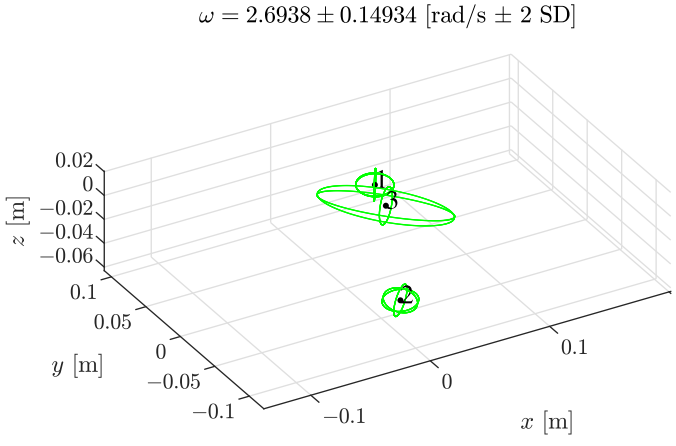
(a) Visualization of the projected flower states after processing 13 images. Flower 3 is initialized for the first time.



(b) Visualization of the CFPM after processing 13 images. Flower 3 is initialized for the first time.

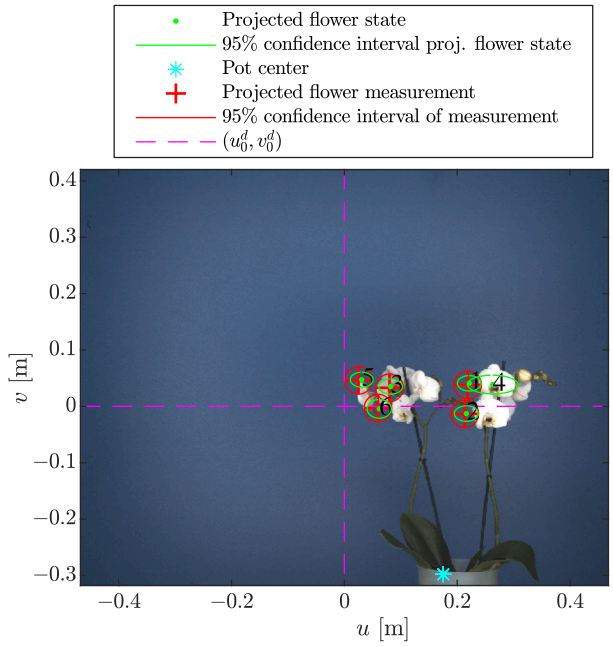


(c) Projected flower states after processing all 20 images.

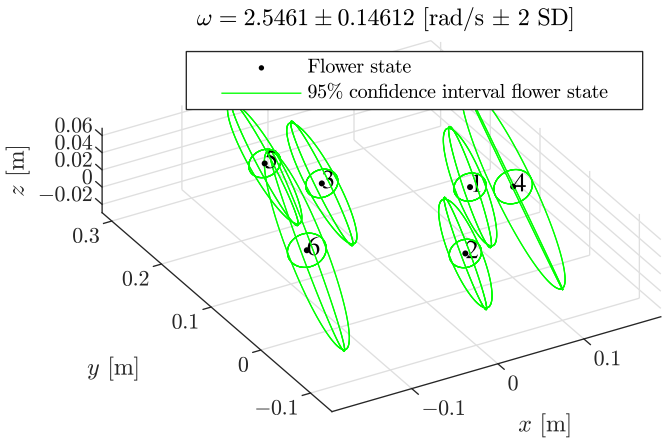


(d) Final estimate of the flower states.

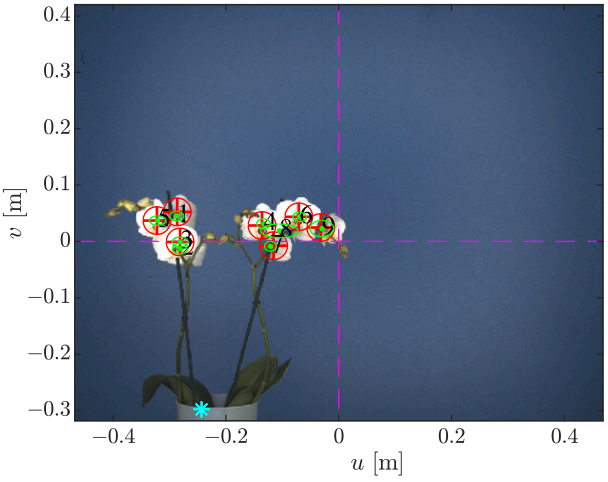
Fig. 8. Visualization of the CFPM (right) and the projected flower states on the images (left) after processing the detections. The plant consists of 3 flowers.



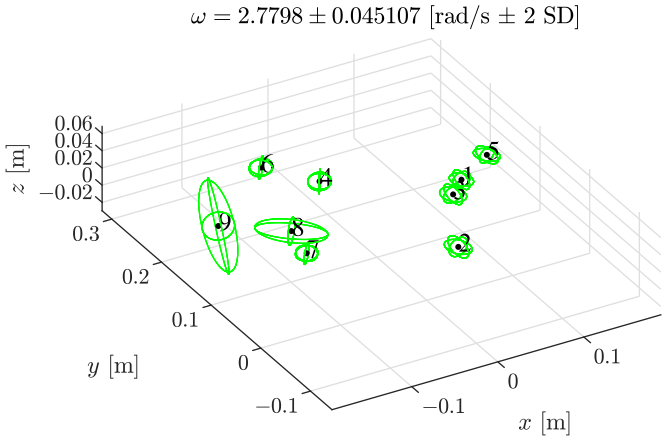
(a) Visualization of the projected flower states after processing 3 images.



(b) Visualization of the CFPM after processing 3 images.



(c) Projected flower states after processing all 20 images.



(d) Final estimate of the flower states.

Fig. 9. Visualization of the CFPM (right) and the projected flower states on the images (left) after processing the detections. The plant consists of 9 flowers.

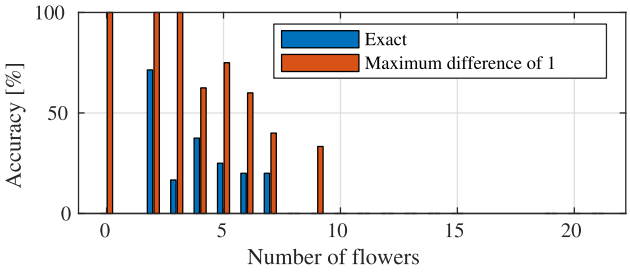
Table 3
Comparative results for the flower quantification when considering the entire dataset.

# Images	Method	Exact	±1 Object
Single	Counting detections	18%	44%
Multiple	Heuristic Approach	32%	58%
Multiple	MHT based on UFPM	37%	70%
Multiple	MHT based on CFPM	61%	92%

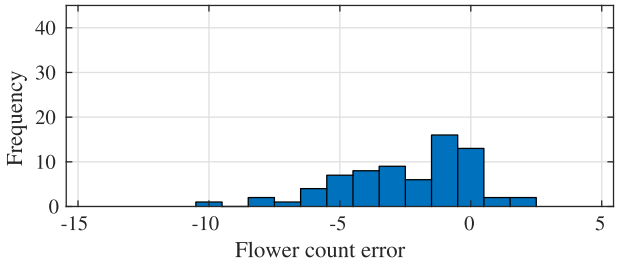
low in relation to the other parameters and are estimated to have an order of magnitude of 10^{-3} [m] and considered equal and independent in each direction. As a result, the process uncertainty $Q_{f_i,k}$ of each flower position is set to

$$Q_{f_i,k} = \mathrm{diag}\left(\begin{bmatrix}\sigma_x^2 & \sigma_y^2 & \sigma_z^2\end{bmatrix}\right). \tag{28}$$

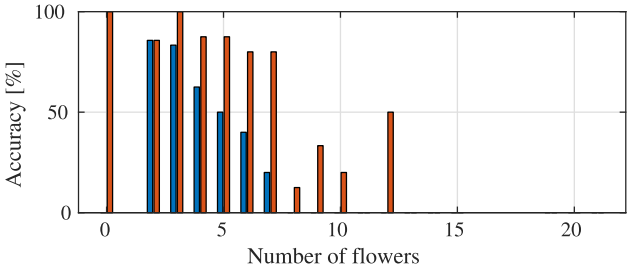
The measurement covariance matrix $S_{f_i,k}$ of flower i is for the CFPM applied in (14) and for the UFPM in (20). The measurement noise of the flower detections is considered equal for each flower and to be bounded by the flower radius in both the u and v direction. As a result, the measurement noise covariance matrix is set to the typical flower radius according to



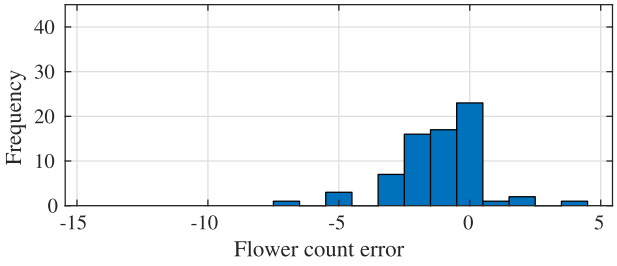
(a) Accuracy for the single image approach.



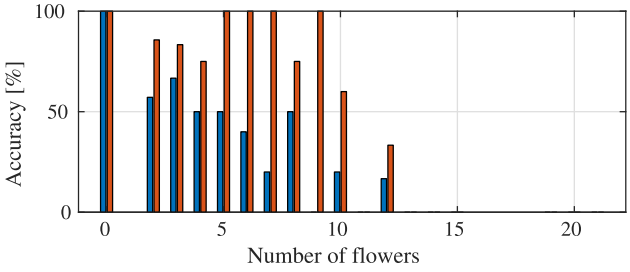
(b) Histogram of the flower estimation error for the single image approach.



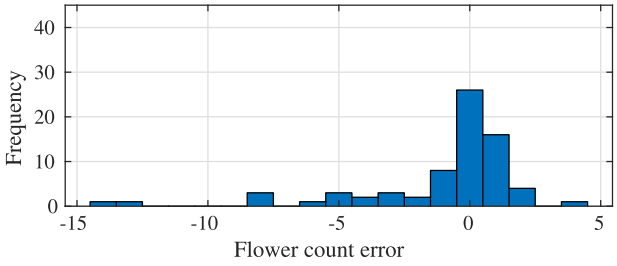
(c) Accuracy for the heuristic approach.



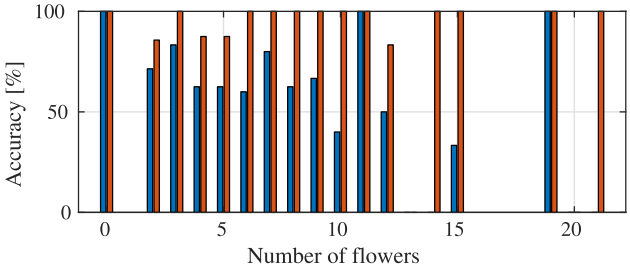
(d) Histogram of the flower estimation error for the heuristic approach.



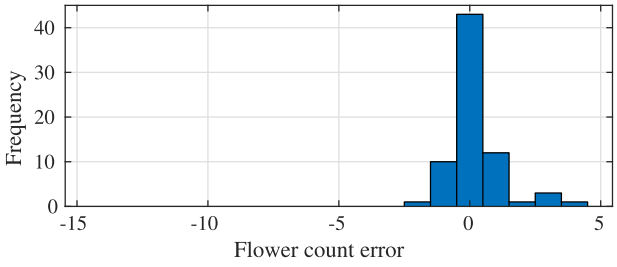
(e) Accuracy for the UFPM based on multiple images.



(f) Histogram of the flower estimation error for the UFPM based on multiple images.

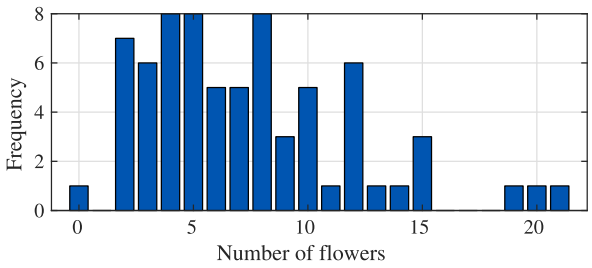


(g) Accuracy for the CFPM based on multiple images.



(h) Histogram of the flower estimation error for the CFPM based on multiple images.

Fig. 10. Results achieved when quantifying the number of flowers in each plant for the methods considered in this work.



(i) Plant distribution of the test set based on the number of flowers.

Fig. 10. (*continued*).

$$S_{f_i,k} = \mathrm{diag}\left(\begin{bmatrix} \sigma_u^2 & \sigma_{u_f}^2 & \sigma_{v_f}^2 \end{bmatrix}\right). \tag{29}$$

4.2. Validation of CFPM and UFPM

For the validation of the methods, the functionality of both the CFPM and the UFPM when running the MHT algorithm is demonstrated and compared. For visualization purposes, a plant with a low number of flowers is shown. Components of the estimated state vector obtained when running the entire MHT algorithm for a plant consisting of three flowers using the independent object models is given in Fig. 7 and Table 2. In this figure, the derived object states of the MAP hypothesis after processing 13 images and after the full set of 20 images is shown on the right side of the figure. To properly assess the relative flower positions, for the UFPM the states are rotated to the moment at which the first flower was detected. The states are visualized in the local frame and the numbers in these figures indicate the identifier of each flower. Due to solving the ambiguities within the MHT approach these identifiers may switch between images due to switching of the MAP hypothesis. Furthermore, the 95% confidence interval of the (projected) flowers is indicated. On the left side of the figure, the recorded images are shown in the center-frame. In this figure, the image center is indicated in magenta and the detections including their 95% confidence interval are indicated in red. Further, the projection of propagated state to measurement space and their corresponding confidence interval are determined according to the observation model and indicated in green. The results of a similar experiment for the CFPM are shown in Fig. 8, while the results of a plant with more flowers, 9 in this case, are given in Fig. 9. For the CFPMs, the rotational velocity is indicated in the corresponding figures.

The upper images of these three figures all indicate a moment where one or more flowers are first seen in the MAP hypothesis. Typically, a relatively large uncertainty in a single direction is observed as the corresponding flower is not observed from multiple viewpoints yet. Based on the projections it is seen that after several update steps the projected state of each flower converges to the desired position, namely the center of the corresponding flower. After processing all evidence of the UFPM as shown in Fig. 7c and Fig. 7d a relative high uncertainty is observed in the u-direction of the image for Flowers 2 and 4. Inspection of the data indicated that this high uncertainty is caused by both flowers not being detected for several frames, leading to an increase in uncertainty in the estimate of the rotational velocity as indicated by Table 2. This results in a false flower as the uncertainty in the estimate of the rotational velocity was too large. Next, caused by large distance between the measurement and the predicted measurement, a better explanation was found by initializing a new object. For the same experiment with the CFPM, visualized in Fig. 8c, the large uncertainty in the u-direction is resolved as the rotational velocity estimation is updated by the detections of the other flowers. For the plant with 9 flowers of Fig. 9 a proper estimation is obtained and even the two sets of flowers which originate from the two branches can be discriminated in the estimated flower states. On both branches in the final state estimate of Fig. 9c, the occlusions are visualized as the states of the occluded flowers are projected on the image while the occluded flowers are hardly visible on the image. Improvements lie in including the estimation of the flower size, such that occlusions could be predicted and associations of detections to occluded flowers could be excluded or have a low probability.

4.3. Greenhouse Dataset

The results when determining the number of flowers for the entire dataset consisting of 71 plants are given in Table 3 and Fig. 10. The correct number of flowers were determined by two persons by visual inspection of the image-sets obtained. In all cases, the persons counted the same number of flowers. The results of the “Single Image”-method are obtained by counting the number of detections for the image where the pot center of the plant was closest to the image center. The choice of the image selection is an arbitrary one. This choice could easily be replaced with another decision and will lead to similar results. The heuristic state of the practice method ranks the detected number of flowers in each image from high to low. Based on the reasoning that in images both occlusions as well as false positives are present, empirical tests indicated that the best results were obtained by considering the image with the third highest number of flowers as the estimated number of flowers. For the MHT algorithm with both the UFPM and the CFPM, the number of flowers was estimated based on the number of flowers present in the MAP hypothesis. Whereas the table quantifies the estimated number of flowers with no deviation and a maximum deviation of 1 flower, the left side of Fig. 10 shows the accuracy for each of the test cases as a function of the number of flowers present in each plant. On the right side of the figure, the counting error of the dataset is demonstrated by means of a histogram. A positive deviation indicates that too many flowers are estimated, no deviation indicates a correct estimation of the number of flowers and a negative deviation indicates an underestimation of the number of flowers. The data indicate that significant improvements (respectively 70% and 92% correct for the UFPM and the CFPM when considering a maximum deviation of 1 flower) are obtained by consistently using multiple viewpoints of each plant compared to a single viewpoint method (44% correct) and the heuristic method (58% correct). As the left side of Fig. 10 indicates, this is especially demonstrated when more flowers are present in a plant. The reason for this is sought in the fact that a larger part of the flowers in a plant are occluded when more flowers are present. This is confirmed by the right side of Fig. 10, as this typically leads to an underestimation of the number of flowers present in the plant. The underestimation of the number of flowers cannot be avoided when using a single image due to occlusions present in Phalaenopsis plants. At the same time, the MHT-based approaches are expected to perform better as well when provided with more accurate and less false positive detections.

Further it is observed that the CFPM representation (92% correct when considering a maximum deviation of 1 flower) is superior to the UFPM (72% correct). It is seen in Fig. 10h that the underestimation of the number of flowers is mostly resolved and that plants with even more flowers are determined correctly. The reason of a slight overestimation in the number of flowers is sought in the flowering of the buds. The detection algorithm can give positive flower detections which are consistent with the model but are not counted in the ground truth. Apart from that, it is concluded that the CFPM suffers less from flower occlusion as the CFPM updates the rotational and translational velocity using flower detections which are correctly associated.

4.4. Future work

For the methods as developed in this work, several options for future work are seen. With respect to the implementation of the current methods in a real time system, which is relevant when sorting the plants based on their number of flowers, the computational efficiency needs to be considered by means of the number of hypotheses. Now, combinatorial associations are initiated as each measurement is separately associated to each modeled object. This includes a set of irrelevant hypotheses. An answer to this is sought in determining better metrics which quantify the plausible flower associations. Considering the entire detection set in the association step at once could for example be beneficial as this creates the opportunity to take neighboring detections and associations into consideration. If for example a detection is found left of another detection at a distance which is expected by the model, the hypothesis which associates the left detection to the expected position of the right object is unlikely and could thus be neglected.

Though improvements in the detection methods are expected given the recent results of deep learning compared to the local binary patterns and AdaBoost learning methods as applied in this work (Kamilaris and Prenafeta-Boldú, 2018; Koirala et al., 2019), the underestimation of the number of flowers cannot be avoided when using a single image due to (full) occlusions as present in Phaleanopsis plants. At the same time, the MHT-based approaches are expected to perform better as well when provided with more accurate and less false positive detections.

Within a robotics context, we see the benefits of this method when tracking for example persons based on partial observations. The observation of for example a hand gesture can indicate some hints about the states of the entire body, even if parts of the body are not detected. This requires however a deeper level of parent–child object relations.

5. Conclusion

To quantify the number of flowers in a plant in the presence of occlusions and detection uncertainty, this work proposed an MHT approach to solve the data association problem given multiple views of the plant. Hereby, the motion uncertainty when obtaining multiple viewpoints is taken into consideration. Under the consideration of a maximum allowed deviation of 1 flower for the quantification of the number of flowers, the CFPM of the MHT target tracker (in which each flower is considered as a part of the plant, 92% accurate as indicated by Table 3) outperformed the UFPM of the MHT (in which each flower is modeled independently, 70% accurate), the heuristic state of the practice counting method (58% accurate) and the single image approach (44% accurate). The most significant differences were found when more flowers and thus occlusions are present, as for the single image approach multiple flowers are not visible at all and for the CFPM the predicted position estimate of unmeasured flowers can still be properly updated based on correct associations of the measured objects. The CFPM-based approach required the MHT to be extended with parent–child relations in the object descriptions to consider the motion model at the parent level and the connections at child level. Whereas in this work a single parent–child relation was sufficient, future research could consider the situation where multiple layers of parent–child relations are required. Furthermore, the use of computational resources of the MHT might be reduced by considering the full detection set of an image. This might help to hypothesize about the uncertain associations only, such that the irrelevant hypotheses could be neglected without sacrificing tracking performance. Despite the room for improvement there still is, given the high accuracy of the MHT based on the CFPM, it was seen that most occlusions and inconsistencies caused by false detection information were solved due to the incorporation of multiple viewpoints.

6. CRediT authorship contribution statement

W. Houtman: Conceptualization, Methodology, Software, Validation, Investigation, Writing - Original Draft, Visualization. **A. Siagkris-Lekkos**: Conceptualization, Methodology, Software, Writing - Review & Editing. **D.J.M. Bos**: Conceptualization, Methodology, Software, Writing - Review & Editing. **B.J.P. van den Heuvel**: Resources, Data Curation. **M. den Boer**: Resources, Supervision. **J. Elfring**: Writing - Review & Editing, Supervision. **M.J.G. van de Molengraft**: Writing - Review & Editing, Supervision.

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.

Appendix A. Nomenclature

An overview of the nomenclature applied in this work can be found in Table 4.

Table 4
Nomenclature as applied in this work. Where necessary, subscripts, superscripts or vectors are indicated to differentiate the variables.

Variable	Meaning
$c()$	cosine-function
D	Distance between camera and image plane
f	Image update frequency
G	Linearized state transition matrix
g	Nonlinear stochastic difference equation
H	Linearized observation matrix
h	Observed state
k	Frame number
I	Identity Matrix
M	Number of flowers in the observed state
N	Number of flowers present in the object state
N_{co}	Number of child objects contained in a hypothesis
N_{po}	Number of parent objects contained in a hypothesis
$\mathscr{M}$	Object behavior model
O	Origin
o	Object
P	State covariance matrix
p	probability
Q	Process covariance matrix
$\mathbf{q}$	Random process noise vector
R	Rotation Matrix
r	Radius
S	Measurement covariance
$\mathbf{s}$	Random measurement noise vector
$s()$	sine-function
t	Time
(u,v)	Image-frame coordinates
(u_0^d,v_0^d)	Image center w.r.t.image origin
(u,v,w)	Center-frame coordinates
$\mathbf{x}$	State vector
(x,y,z)	Plant-frame coordinates
Z	Measurement set

(continued on next page)

Table 4 (*continued*)

$\mathbf{z}$	Observation vector
α	Angle between the w-axis and the sight line
β	Angle between the u-axis and the sight line
θ	Plant rotation
ι	Identifier
Λ	Covariance matrix along sight line
σ	Standard deviation
ω	Rotational velocity of the pot center

Acronym	**Meaning**
CFPM	Connected-Flower Plant Model
EKF	Extended KF
GNN	Global Nearest Neighbor
JPDAF	Joint Probabilistic Data Association Filter
KF	Kalman Filter
MAP	Maximum A Posteriori
MHT	Multiple Hypothesis Tracker
MTT	Multiple Target Tracking
SD	Standard Deviation
UFPM	Unconnected-Flower Plant Model

Subscript	**Meaning**
0	Initial
a	Object index
ap	Parent object index
ac	Child object index
c	Center
d	Detection
f	Flower
i	Flower identifier
j	Measurement index
l	Local frame
m	Dimension of the measurement vector
max	Maximum
n	Dimension of the state vector
P	Plant
p	Projected
pc	Pot center
pf	projected flower
sl	Sight line
T	Transpose

References

Bairwa, N., Agrawal, N.K., 2014. Counting of flowers using image processing. Int. J. Eng. Res. Technol. (IJERT) 03 (09), 775–779. URL https://www.ijert.org/counting-of-flowers-using-image-processing.

De Laet, T., 2010. Rigorously Bayesian Multitarget Tracking and Localization (rigoureus bayesiaans detecteren en volgen van meerdere objecten). https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.296.8829&rep=rep1&type=pdf.

de Waard, H., 2008. A New Approach to Distributed Data Fusion. Ph.D. thesis. Universiteit van Amsterdam.

Dubey, S.R., Jalal, A.S., 2015. Application of image processing in fruit and vegetable analysis: a review. J. Intell. Syst. 24 (4), 405–424. https://doi.org/10.1515/jisys-2014-0079.

Elfring, J., 2021. WIRE, accessed: 2021–01-28. http://wiki.ros.org/wire.

Elfring, J., van den Dries, S., van de Molengraft, M., Steinbuch, M., 2013. Semantic world modeling using probabilistic multiple hypothesis anchoring. Robot. Auton. Syst. 61 (2), 95–105. https://doi.org/10.1016/j.robot.2012.11.005.

Hamidinekoo, A., Garzón-Martínez, G.A., Ghahremani, M., Corke, F.M.K., Zwiggelaar, R., Doonan, J.H., Lu, C., 2020. DeepPod: a convolutional neural network based quantification of fruit number in Arabidopsis. GigaScience 9 (3), giaa012. https://doi.org/10.1093/gigascience/giaa012.

Harmsen, S.R., Koenderink, N.J., 2009. Multi-target tracking for flower counting using adaptive motion models. Comput. Electron. Agric. 65 (1), 7–18. https://doi.org/10.1016/j.compag.2008.07.004.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Koirala, A., Walsh, K.B., Wang, Z., McCarthy, C., 2019. Deep learning – method overview and review of use for fruit detection and yield estimation. Comput. Electron. Agric. 162, 219–234. https://doi.org/10.1016/j.compag.2019.04.017.

Koporec, G., Pers, J., 2019. Deep learning performance in the presence of significant occlusions - an intelligent household refrigerator case. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE Computer Society, Los Alamitos, CA, USA, pp. 2532–2540. https://doi.org/10.1109/ICCVW.2019.00310.

Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., Kim, T.-K., 2021. Multiple object tracking: a literature review. Artif. Intell. 293, 103448. https://doi.org/10.1016/j.artint.2020.103448.

Mahmud, M.S.A., Abidin, M.S.Z., Emmanuel, A.A., Hasan, H.S., 2020. Robotics and automation in agriculture: present and future applications. Appl. Modell. Simulat. 4, 130–140. URL http://arqiipubl.com/ojs/index.php/AMS_Journal/article/view/130.

Moonrinta, J., Chaivivatrakul, S., Dailey, M.N., Ekpanyapong, M., 2010. Fruit detection, tracking, and 3d reconstruction for crop mapping and yield estimation. In: 2010 11th International Conference on Control Automation Robotics Vision, pp. 1181–1186. https://doi.org/10.1109/ICARCV.2010.5707436.

Naranjo-Torres, J., Mora, M., Hernández-García, R., Barrientos, R.J., Fredes, C., Valenzuela, A., 2020. A review of convolutional neural network applied to fruit image processing. Appl. Sci. 10 (10) https://doi.org/10.3390/app10103443.

Puttemans, S., Goedemé, T., 2015. Visual detection and species classification of orchid flowers. In: 2015 14th IAPR International Conference on Machine Vision Applications (MVA), pp. 505–509. https://doi.org/10.1109/MVA.2015.7153241.

Robin, C., Lacroix, S., 2016. Multi-robot target detection and tracking: taxonomy and survey. Auton. Robots 40 (4), 729–760. https://doi.org/10.1007/s10514-015-9491-7.

Sethy, P.K., Routray, B., Behera, S.K., 2019. Detection and counting of marigold flower using image processing technique. In: Biswas, U., Banerjee, A., Pal, S., Biswas, A., Sarkar, D., Haldar, S. (Eds.), Advances in Computer, Communication and Control. Springer Singapore, Singapore, pp. 87–93. https://doi.org/10.1007/978-981-13-3122-0_9.

Song, Y., Glasbey, C., Horgan, G., Polder, G., Dieleman, J., van der Heijden, G., 2014. Automatic fruit recognition and counting from multiple images. Biosyst. Eng. 118, 203–215. https://doi.org/10.1016/j.biosystemseng.2013.12.008.

Spong, M., Hutchinson, S., Vidyasagar, M., 2005. Robot Model. Contr. Wiley. URL https://books.google.nl/books?id=P4Y4xQEACAAJ.

Stein, M., Bargoti, S., Underwood, J., 2016. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. Sensors 16 (11). https://doi.org/10.3390/s16111915.

Syal, A., Garg, D., Sharma, S., 2014. Apple fruit detection and counting using computer vision techniques. In: IEEE International Conference on Computational Intelligence and Computing Research, pp. 1–6. https://doi.org/10.1109/ICCIC.2014.7238364.

Tang, Y., Chen, M., Wang, C., Luo, L., Li, J., Lian, G., Zou, X., 2020. Recognition and localization methods for vision-based fruit picking robots: a review. Front. Plant Sci. 11, 510. https://doi.org/10.3389/fpls.2020.00510.

Thrun, S., Burgard, W., Fox, D., 2005. Probabilistic Robotics (Intelligent Robotics and Autonomous Agents). The MIT Press.

Tripathi, M.K., Maktedar, D.D., 2020. A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: a survey. Inform. Process. Agric. 7 (2), 183–203. https://doi.org/10.1016/j.inpa.2019.07.003.

Wäldchen, J., Mäder, P., 2018. Plant species identification using computer vision techniques: a systematic literature review. Arch. Comput. Methods Eng. 25 (2), 507–543. https://doi.org/10.1007/s11831-016-9206-z.

Wang, Q., Nuske, S., Bergerman, M., Singh, S., 2013. Automated crop yield estimation for apple orchards. In: Desai, J.P., Dudek, G., Khatib, O., Kumar, V. (Eds.), Experimental Robotics: The 13th International Symposium on Experimental Robotics. Springer International Publishing, Heidelberg, pp. 745–758. https://doi.org/10.1007/978-3-319-00065-7_50.

Wang, J., Zhang, Z., Xie, C., Zhu, J., Xie, L., Yuille, A., 2017. Detecting Semantic Parts on Partially Occluded Objects. BMVA Press, pp. 73.1–73.13. https://doi.org/10.5244/C.31.73.

Wei, J., Zhijie, Q., Bo, X., Dean, Z., 2018. A nighttime image enhancement method based on retinex and guided filter for object recognition of apple harvesting robot. Int. J. Adv. Robot. Syst. 15 (1) https://doi.org/10.1177/1729881417753871.

Welch, G., Bishop, G., 2006. An Introduction to the Kalman Filter. Tech. rep., University of North Carolina at Chapel Hill. URL http://www.cs.unc.edu/~welch/kalman/kalmanIntro.html.

Yang, C., Chong, P.H.J., Lam, P., 2021. An occlusion handling evaluation criterion for deep learning object segmentation. J. Phys: Conf. Ser. 1880 (1), 012008. https://doi.org/10.1088/1742-6596/1880/1/012008.