A Model for Pollen Measurement Using Video Monitoring of Honey Bees

Article *in* Sensing and Imaging An International Journal · December 2018

DOI: 10.1007/s11220-017-0185-4

3 authors, including:

Cheng Yang
Auckland University of Technology
6 PUBLICATIONS **8** CITATIONS
SEE PROFILE

John Collins
Auckland University of Technology
13 PUBLICATIONS **38** CITATIONS
SEE PROFILE

Sens Imaging (2018) 19:2
https://doi.org/10.1007/s11220-017-0185-4

ORIGINAL PAPER

A Model for Pollen Measurement Using Video Monitoring of Honey Bees

Cheng Yang[1] · **John Collins**[1] · **Mark Beckerleg**[1]

Received: 28 June 2017 / Revised: 14 September 2017

Abstract This paper introduces a novel approach using video monitoring to automatically count pollen sacs on honey bees returning to a beehive. This allows beekeepers to have an estimate of the food storage situation in a beehive without having to visit and open the hive. Pollen sac detection is a challenge because the full sacs are very small and they often have a similar colour to the bee's body. The first step is to detect bees on the video and track the bees for pollen counting. The bees are detected using their colour and motion, and tracked using a Kalman filter and the Hough transform. The second step is to detect pollen sacs on each individual detected bee image. On each individual bee image, a bee blob is created and image moments are used to remove an elliptical shape from the blob, which includes the bee's main body parts (head, abdomen and tail). The pollen sacs are then detected in the remainder of the blob using colour thresholding. The outcome includes pollen blobs and other blobs that are similar to pollen blobs (non-pollen blobs). The pollen sacs have two features that can be used for object discrimination. The receiver operating characteristic is used to calculate the threshold for pollen discrimination using these two features. Finally, the bee tracking model is combined with the pollen sac detection to count the number of bees carrying pollen on the monitoring video.

✉ Cheng Yang
royang@aut.co.nz

John Collins
jcollins@aut.ac.nz

Mark Beckerleg
mbeckerl@aut.ac.nz

[1] School of Engineering, Auckland University of Technology, 34 St. Paul Street, City Campus, Auckland, New Zealand

Published online: 29 December 2017

Keywords Pollen detection · Pollen measurement · Kalman filter · Hough transform · Receiver operating characteristic

1 Introduction

In the honey industry, beekeepers focus on the bees' storage and production of honey. Beekeepers often need to open beehives to check the storage situation. This costs time and labor. This paper introduces a video processing model to automatically count bees with full pollen sacs using a monitoring video located outside the beehive. Computer vision techniques are used to identify and count the pollen sacs. Pollen sac detection is a challenge, because the pollen sacs are very small and they have similar colour to the bees' bodies. In the research literature, there appear to be no publications concerned with pollen detection on bee's bodies.

Other researchers have tried to detect objects which are as small as the pollen sacs. The aim of small object detection is to detect a special small region on a much larger image. Although the object's area may be too small to be captured directly, papers have reported detecting small object regions using various methods. Desai et al. [1] combined the methods of wavelet and gradient to detect small objects. Wei [2] detected objects of small size and fast motion in Infra-Red (IR) sequences through manipulations in the temporal, spectral and spatial domains. This method worked even though the IR images were noisy. Baojun et al. [3] also used IR images for detection of small maritime objects. They used wavelet transforms and the histogram method for the detection. Vard et al. [4] reported an active contour model to detect small objects from cluttered and textured backgrounds. Qi et al. [5] presented an algorithm for small surface object detection using depth-aware analysis of image features by the atmospheric scattering model on maritime scenes video. Hsieh et al. [6] used a watershed-based transformation to detect small objects.

These papers successfully detect the small objects, but the small object was obvious on the background in most of these papers. Both [2, 3] obtained their results using infrared images, where the object was a white point on the background before the detection. The small points in [4] were also light points which were obvious against the dark background around them. The backgrounds in [1, 5, 6] are simple sky or ocean scenes, and the detected airplanes or ships were clearly visible in those situations.

The above methods are difficult to apply to more complicated situations. Pollen sacs, as an example, also appear very small on the video. In addition, their colour is often very similar to bees' bodies. The video frames may also include motion blur. Although research concerning pollen detection has not been found, research has been reported concerning detection of honey bees on an image or video. Campbell et al. [7] used motion models to detect and track the flight of bees in the uncontrolled outdoor environment of an apiary. The result is affected by the bee shadows, even on a single colour background. Kimura et al. [8] reported a method to detect and track bees on a flat surface. They employed a background subtraction method to detect bees. However, because the moving was only a few pixels, if the bee moved quickly then the tracking failed. 3D stereo vision systems have also been

used to detect and track bees [9, 10]. The 3D camera recorded video with two views. One view was an intensity video that was used to detect bee flight using background subtraction. The other view was a colour image, from which the colour histogram method was used to detect bees. Finally, a Kalman filter was used to track bees on the video. This 3D system required substantial computation time.

The aim of this research is to detect and count pollen sacs on the bee monitoring video. The video is recorded by a camera mounted on a beehive. Figure 1 shows the beehive and the camera set up situation. The camera is attached to the front wall of the beehive, facing down, about 30 cm above the entrance. The video is 1920 * 1080 resolution at a rate of 50 frames per second. The camera is programmed to turn on every half hour. Once it is on, the camera records the video for 2 min, and then turns itself off automatically. After the next half hour, it repeats the operation. In this way, video is recorded at different times of the day.

A single green colour board is placed on the ground in front of the entrance to the beehive, to simplify the image background. This is necessary, because the natural background includes withered grass with a similar colour to the bees and this interferes with the image segmentation. The final view of the videos is shown in Fig. 2. It is difficult to record pollen on the video even in the flowering season. Videos including pollen sacs are chosen manually for analysis.

The videos are transferred to a computer, and the processing is implemented using MATLAB with the computer vision toolbox. MATLAB is used to develop a model that can automatically detect and count pollen sacs on the video.

Fig. 1 The beehive and camera setup. Note: the white board on the bottom has changed to green in the final setting (Color figure online)

Fig. 2 The view from the monitoring camera (Color figure online)

The first step of this research is to detect and track bees on the video. This simplifies the pollen detection, because the pollen only needs to be detected from individual bee images. The bees are detected using colour thresholding and motion detection with a Gaussian mixture model. Bees are flying in front of the beehive and their shadows are also moving. The motion detection captures both the bee and shadow motion. However, colour can be used to remove shadows and obtain clearer bee shapes. The Kalman filter is then used to track bees on the video. The Hungarian assignment method is combined with the Kalman filter to assign the predicted and measured positions for multiple bee tracking. The resolution of the video is 1080p, so bee blobs are big. A problem occurs when bees fly across each other in the video image, creating a merged bee situation. The Kalman filter cannot successfully track each individual bee in the merged situation. This problem is solved using the Hough transform. This technique uses the shape of individual bees to estimate each bee position in the merged blob. The bee tracking can be used to count the number of bees appearing in the video, so that the number of pollen sacs can also be counted later.

The second step is pollen detection and measurement. The individual bee images and blobs are cropped out of the video frame. Then the individual bee blobs are analysed using image moments to remove the main bodies (including heads, abdomens and tails) of bees. Colour features are used to identify pollen sacs (if they are present) to separate them from other bee features such as wings and legs. The result is a collection of candidate blobs that include pollen sac blobs (pollen blobs) and other blobs that are similar to pollen sacs (non-pollen blobs). The receiver operating characteristic (ROC) technique is used to calculate the thresholds used to discriminate the pollen blobs from the non-pollen blobs. Two features of the pollen blobs are collected for ROC analysis to find the feature thresholds for pollen sac detection.

The last step is to combine the bee tracking and pollen detection on the video to count number of bees that are carrying pollen sacs. Each bee appears on the video in a number of frames. It is necessary to track the bee to avoid counting the same pollen sac multiple times. In addition, in some frames the pollen sac may not be visible. The accuracy of the pollen sac count is improved by using the number of observations of pollen sacs on each individual bee as it is tracked in the video, until it is disappeared.

The rest of this paper is organized as follows. Bee detection and tracking is described in Sect. 2, Sect. 3 describes the pollen detection and discrimination, Sect. 4 introduces the full model of the pollen measurement on the video, the experiment results are presented in Sect. 5, and conclusions are given in Sect. 6.

2 Bee Detection and Tracking

2.1 Bee Detection

The bee detection method uses a combination of motion detection and colour thresholding. Because the colour of bee bodies includes black, orange, yellow, red and white (some pollen sacs have red or white colour), the background of the video is simplified by placing a green board on the ground below the hive entrance. The Gaussian Mixture Model (GMM) [11] is utilized to detect moving bees and ignores the unchanging uniform background. However, this also detects bees' shadows, which affects the result. The colour thresholding uses the Hue-Saturation-Value (HSV) colour space [12] to detect the orange and black colour of bees. This does not detect the shadow, but may lose some pixels and detect other things that include the same colour as bees, such as pollen dust. The combination of motion and colour methods removes their disadvantages to obtain an accurate shape for the bees. The theory has been reported in [13].

After this processing, the final binary image shows the bee blobs. The bee blobs can be analysed to calculate various parameters of the blob and draw rectangular bounding boxes (bbox) around each bee. The bounding box is used to crop the bee image from the background.

2.2 Tracking Method

The video resolution is suitable for bee detection, but it is difficult to track bees because they move a significant distance between successive frames in the video, even though the video records 50 frames per second. A Kalman filter is used to track bees.

The discrete Kalman filter at frame number k is described below.

$$\text{Prediction:} \quad s_k^- = A\left[s_{k-1}^+, s_{k-2}^+\right]^T \tag{1}$$

$$P_k = AQ_{k-1}A^T + Q_0 \tag{2}$$

$$\text{Measurement:} \quad m_k = Hs_k \tag{3}$$

$$\text{Kalman gain:} \quad K_k = P_k(P_k + R_k)^{-1} \tag{4}$$

$$\text{Correction:} \quad s_k^+ = s_k^- + K_k\left(m_k - s_k^-\right) \tag{5}$$

$$Q_k = (I - K_k)P_k \tag{6}$$

In Eqs. (1)–(6), s_k is the actual state that the detection is measuring (the x and y coordinates of the bee). s_k^- is the predicted state, and s_k^+ is the corrected state after combining the measurement m_k with s_k^-. P_k is the covariance matrix of the prediction s_k^-, R_k is the covariance matrix of the measurement m_k, and Q_k is the covariance matrix of the corrected state s_k^+. The matrix H is the measurement model matrix, which is the identity matrix in this case. The correction is controlled by the Kalman gain of K_k which is calculated from P_k and R_k. In this paper, for individual bees, the measurement (m_k) (bee detection) is more reliable than the prediction (s_k^-), so the covariance P_k is greater than R_k. Then the estimated state s_k^+ is closer to the measurement than to the prediction.

In this research, Q_0 is the variance due to the bee velocity being variable. It adds to the covariance of the prediction due to previous calculations. Q_0 is not only added once at the beginning of tracking, it is added in each step of the Kalman filter calculation.

To apply the Kalman filter formulas, the state of a bee is its position (x, y), and the prediction of the Kalman filter tracking model in frame k is:

$$x_k^- = 2x_{k-1}^+ - x_{k-2}^+ \tag{7}$$

$$y_k^- = 2y_{k-1}^+ - y_{k-2}^+ \tag{8}$$

The prediction of the Kalman filter is:

$$s_k^- = A\left[s_{k-1}^+, s_{k-2}^+\right]^T \tag{9}$$

and the predicted state vector is:

$$s_k^- = \left[x_k^-, y_k^-\right]^T \tag{10}$$

The corrected state vector in the frame k is:

$$s_k^+ = \left[x_k^+, y_k^+\right]^T \tag{11}$$

The transition matrix is defined as:

$$A = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 \end{bmatrix} \tag{12}$$

The measurement of the Kalman filter is:

$$m_k = Hs_k \tag{13}$$

The actual state vector is:

$$s_k = [x_k, y_k]^T \tag{14}$$

The measurement transition matrix is:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{15}$$

The measurement vector is:

$$m_k = [x_k, y_k]^T \tag{16}$$

The covariance matrix of the measurement:

$$R_k = \begin{bmatrix} r_{11k} & 0 \\ 0 & r_{22k} \end{bmatrix} \tag{17}$$

The covariance matrix of correction:

$$Q_k = \begin{bmatrix} q_{11k} & 0 & 0 & 0 \\ 0 & q_{22k} & 0 & 0 \\ 0 & 0 & q_{11k-1} & 0 \\ 0 & 0 & 0 & q_{22k-1} \end{bmatrix} \tag{18}$$

After setting up the Kalman filter for each bee on the video, the Hungarian assignment method [14] is used for multiple bee tracking. This method assigns predictions to detections by minimizing the sum of the distances between the assigned predictions and detections. As a constraint, a prediction will not be assigned to a detection when the separation distance is more than 100 pixels. This will usually result in some unassigned predictions and detections. When a detection is unassigned, it is usually a bee appearing in the video for the first time. In addition, unassigned predictions can be caused by a bee disappearing from the video or when bees fly across each other (merge) in the video. In this case, bee detection produces one big blob rather than separated individual (single) bee blobs. The merging occurrence is identified by checking whether an assigned detection is near the unassigned prediction or not.

The covariance values are estimated from the errors between the system calculation of the bee positions in the model and the actual bee positions as calculated manually. It is assumed the x and y values are independent of each other, so both the covariance matrices are diagonal.

For individual (single) bee tracking, 172 single bee images were used to estimate errors and calculate the variances. The measurement covariance matrix was found to be:

$$R_k = \begin{bmatrix} 13 & 0 \\ 0 & 10 \end{bmatrix} \tag{19}$$

In addition, the covariance matrix Q_0 was calculated from bee positions on the images.

$$Q_0 = \begin{bmatrix} 124 & 0 \\ 0 & 78 \end{bmatrix} \tag{20}$$

These two covariance matrices are assumed to be constant during single bee tracking. However, because bees can merge with each other in the image, the prediction and measurement covariance matrices in the merged tracking are different from the single tracking. 91 merged bee images were utilised to get the following matrices for the merged situation:

$$R_k = \begin{bmatrix} 419 & 0 \\ 0 & 717 \end{bmatrix} \tag{21}$$

$$Q_0 = \begin{bmatrix} 435 & 0 \\ 0 & 825 \end{bmatrix} \tag{22}$$

These two covariance matrices are also assumed to be constant during merged bee tracking.

2.3 Hough Transform for Merging

A difficulty occurs when two bees pass across each other in the image. This is merging of the individual bee images. The merging may also involve more than two bees. After the merging has been identified, the Hough transform method is applied to estimate the location of each bee in the merged blob. The feature for the Hough transform is the single bee's shape. The concept is similar to circle detection [15].

Before a bee merges with other bees, its shape can be determined from its blob image. If the bee merges in the next frame, the merged blob outline partly shows the single bee's shape. Suppose the edge image (I_s) of the single bee has width $\boldsymbol{I}$ and height $\boldsymbol{J}$. The image pixel at i,j has the value:

$$I_s(i,j) = \begin{cases} 1, & \textit{on the edge} \\ 0, & \textit{not on the edge} \end{cases}, \tag{23}$$

where $1 \leq i \leq I$ and $1 \leq j \leq J$. The centre of the image is at (i_c, j_c) which is also the shape centre. The pixels with value "1" define the edge curve on the image I_s. The set of pixels (**S**) on the edge of the single bee blob is:

$$\mathrm{S} = \{(i,j) | I_s(i,j) = 1\} \tag{24}$$

It is assumed that the set of edge pixels has N members. Each member of the set **S** is denoted by $s_n = (i_n, j_n)$ representing the coordinates of edge pixel number $(n \in [1, N])$.

In the edge image, the origin is at the top left corner. To detect a single bee in the merged blob, it is convenient to shift the origin to the centre of the edge image. The transformation is:

$$c_n = [u_n, v_n] = [i_n - i_c, j_n - j_c], \tag{25}$$

so that the transformed edge pixel set (**C**) is:

$$\mathrm{C} = \{c_n | n \epsilon [1, N]\} \tag{26}$$

When the bee merges with other bees, the merged blob can be used to produce the merged shape (or edge). Suppose, the edge image of the merged blob (I_m) has width X and height Y, and its pixels are:

$$I_m(x, y) = \begin{cases} 1, & \textit{on the edge} \\ 0, & \textit{not on the edge} \end{cases}, \tag{27}$$

where $1 \leq x \leq X$ and $1 \leq y \leq Y$. The pixels with value "1" are on the edge of the blob on the image I_m. The set of pixels (M) on the edge of the merged blob is:

$$\mathrm{M} = \{(x, y) | I_m(x, y) = 1\} \tag{28}$$

This set of pixels has K members. Each member is denoted by $m_k = (x_k, y_k)$ representing the coordinates of pixel number $(k \in [1, K])$.

In the Hough transform technique, the edge pixels of the single bee blob are drawn on the merged blob edge image, with the centre of the single bee blob positioned on each pixel on the edge of the merged blob. Therefore, if the merged blob edge has $\boldsymbol{K}$ pixels, the single blob edge is drawn $\boldsymbol{K}$ times on the merged blob edge image. The coordinates of the resulting pixels drawn on the image are

$$t_{nk} = [u_n + x_k, v_n + y_k] \tag{29}$$

This transform set is known as T, where:

$$T = \{t_{nk} | n \epsilon [1, N], k \epsilon [1, K]\} \tag{30}$$

These are the Hough Transform pixels. Some of the pixels t_{nk} are in the same position as each other.

Each pixel in the merged blob edge image (I_m) may match none, one or several of the Hough Transform pixels (T). This creates a voting map ($V(\mathrm{x}, \mathrm{y})$), whose size is the same as the size of the merged blob edge image. Each element of the voting map array corresponds to a pixel of the merged edge image. These elements are bins that record the number of Hough Transform pixels located at position (x, y). For (x, y) inside the merged blob, $V(\mathrm{x}, \mathrm{y})$ is the number of t_{nk} values equal to (x, y). These points are $x \epsilon [1, X], y \epsilon [1, Y]$. $V(\mathrm{x}, \mathrm{y})$ is also called the voting value (or voting map).

The voting map indicates the possible positions of the single bee on the merged image. If the single bee blob edge fits part of the merged blob edge, there is expected to be a corresponding element of voting map that has a peak value. This peak point is a candidate for the single bee's position in the merged blob.

However, the motion of the bee can change its orientation by up to $\pm 10^{\circ}$ between successive video frames. If the bee changes orientation, the voting map may produce a lower valued peak point. Conversely, the correct fitting orientation will cause the voting map to get a higher peak point. The rotation of the bee's blob is described as:

$$\alpha = \{\alpha_l | l \epsilon 1, 2, \ldots, L\}, \tag{31}$$

where $L = 21$ and the

$$[\alpha_1, \ldots \alpha_{21}] = [-10^{\circ}, -9^{\circ}, \ldots, 0^{\circ}, \ldots, 9^{\circ}, \ldots, 10^{\circ}] \tag{32}$$

are the angles which are used to rotate the single bee blob shape from its original orientation. The plus sign indicates an anticlockwise rotation, and the minus sign indicates a clockwise rotation. If the single bee blob edge is rotated to α_l, the edge pixel coordinates relative to the centre of the single bee blob are:

$$c_{n\alpha_l} = [u_n \times \cos \alpha_l + v_n \times \sin \alpha_l, v_n \times \cos \alpha_l - u_n \times \sin \alpha_l] \tag{33}$$

Each rotated single bee blob edge curve can be transformed to each voting map for an α_l rotation. In this case, there are 21 different voting maps corresponding to the 21 different rotations. The different voting maps display different valued peak points. If a rotation makes the single bee edge fit the merged edge, a peak point in the corresponding voting map is the highest, with other peak points being much lower. If the rotation does not fit the merged edge, there may be two or three peak points with a similar, but lower, value.

However, the fitting rotation may not be the highest peak value. This is because the single bee shape sometimes fits another part of the merged shape rather than the correct part, so that the wrong position also includes the high peak point. If the peak point is near the correct position, the region around this point includes many higher voting value points. This is described as the correct region. Conversely, the region around the wrong position has some low value points, and this is known as the incorrect region, so that this peak point is sharp and isolated.

Therefore, the sum of the region values around the top points can be used to find the correct region. In each of the 21 voting maps, all the values at voting points within $\pm$ 5 pixels of the top point are added together. It is assumed the top point in the voting map for orientation index $\boldsymbol{l}$ is $V(x_p, y_p)_l$, and the total value of the region around this top point is:

$$r_l = \sum_{x=x_p-5}^{x_p+5} \sum_{y=y_p-5}^{y_p+5} V_l(x, y) \tag{34}$$

where $l \in [1, 21]$. The top value of the 21 region total values produces the correct region that includes the actual position of the bee.

The top point in the correct region might be expected to be the candidate position. However, this point is only near the actual position, because of the shape noise. The shape noise comes from the segmentation. It is difficult to get accurate shape detail when performing segmentation, but the detection can be improved by a

weighted average calculation near the correct region. The correct region itself is too small to be used to modify the position calculation. The region is extended to an area of ± 20 pixels from the peak point, and a weighted average of the voting map is calculated as below:

$$w_x = \frac{\sum_{x=x_p-20}^{x_p+20} \sum_{y=y_p-20}^{y_p+20} xV(x,y)}{\sum_{x=x_p-20}^{x_p+20} \sum_{y=y_p-20}^{y_p+20} V(x,y)} \tag{35}$$

$$w_y = \frac{\sum_{x=x_p-20}^{x_p+20} \sum_{y=y_p-20}^{y_p+20} yV(x,y)}{\sum_{x=x_p-20}^{x_p+20} \sum_{y=y_p-20}^{y_p+20} V(x,y)} \tag{36}$$

In the formula above, the point at (w_x, w_y) replaces the peak point as the candidate position of the bee.

This is the final detection of the bee on the merged image. Other merged bees are detected in the same way. In the following merged frame, the fitting rotation still uses the original orientation, and the single bee detection that is just before the first merged frame is still used to detect the merged bee. This method is used until the two bees separate.

The Hough transform improves the Kalman filter for merged bee tracking. The whole tracking model includes the Kalman filter and Hough transform.

3 Pollen Detection and Recognition

3.1 Pollen Detection

From Sect. 2.1, the individual bees are detected from the frames of the video. The blobs of the bees are shown on the binary image of each frame. The individual bee images are cropped from the video frames. The blobs are also cropped from the binary images.

Figure 3 shows an example of detection of an individual bee. Image (a) is a detected individual bee, cropped from the monitoring video. A pollen sac is clearly visible as a light yellowish blob to the left of the centre of the image. The size of this image is 156 × 97 pixels (height × width). The bee's detected binary blob is shown in image (b). The main body of the bee includes orange colour which is similar to the pollen sac colour, so before pollen detection, the main body (head, abdomen and tail) is removed by using image moments which analyse the bee's blob.

Image moments are used to analyse the bee blob binary image shown in Fig. 3b. The outputs of the method are the elliptical parameters of the blob: centre position (x_c, y_c), orientation (θ), height (h), width (w), major axis ($2a$) and minor axis ($2b$). These parameters are used to construct an ellipse, which is

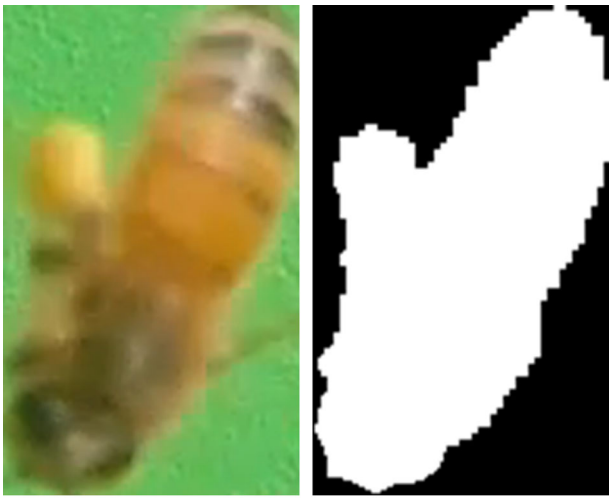

Fig. 3 An example of an individual bee detection. **a** The cropped image of the individual bee. **b** The cropped blob as a binary image (Color figure online)

$$\frac{((x-x_c)\cos\ +(y-y_c)\sin)^2}{a^2}+\frac{((x-x_c)\sin\theta-(y-y_c)\cos)^2}{b^2}=1 \qquad (37)$$

The ellipse is smaller than the bee blob. It does not include projections outside the main body of the bee. This is shown in Fig. 4b. Projections outside the ellipse are discarded. Figure 4c is the smaller blob. This blob is analysed again by using image moments to calculate another ellipse (Fig. 4d). Again, projections outside the ellipse are discarded to produce a smaller blob in Fig. 4e. This process repeats until no projections are left. The result is an elliptical blob in Fig. 4f. Compared to the original bee image of Fig. 4g, the ellipse can be used to mark the main body of the bee.

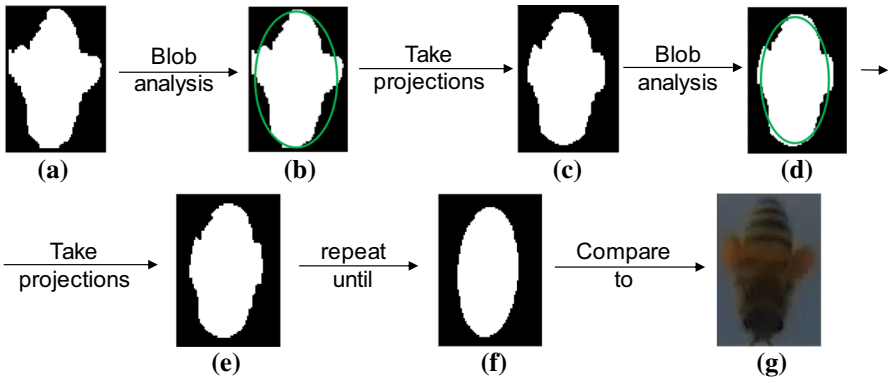

Fig. 4 An example of bee's main body blob detection

Image Fig. 5a is the main body blob of the bee in Fig. 3 which is detected using the method above. The main body blob is used to remove the main body from the bee image. The result is displayed in image Fig. 5b.

After the main body removal, the pollen sac from Fig. 5b can be detected using colour thresholding. Pollen sacs can have white, yellow and orange colours. In the example of Fig. 6, the pollen sac is yellow. The result of colour thresholding is shown in this figure. It can be seen that not only the pollen sac is found, but also there are other non-pollen blobs including parts of the tail, thorax and other small dust particles. Therefore, this detection result is not accurate enough. In addition, if a bee does not have a full pollen sac, the above method still detects some non-pollen blobs.

One monitoring video was chosen to refine the pollen detection method. In the first 400 frames of the video, 84 bees are detected as whole bodies using the bee detection method. Only 10 of these bees actually have pollen sacs. There are 276 individual bee images of these 10 bees in the 400 frames. Not all of these images have visible pollen sacs, because sometimes the pollen sacs are obscured by the bee's body. In addition, 74 bees without full pollen sacs are detected as whole bodies in the 400 frames. These 74 bees are detected and cropped into 1320 individual bee images. All of these images do not have pollen sacs. In total, the pollen detection method detects 4255 blobs from 1596 (276 + 1320) individual bee images. Of these blobs, 273 are pollen blobs; and 3982 are non-pollen blobs. It is necessary to distinguish the pollen blobs and non-pollen blobs.

Fig. 5 The main body (head and abdomen) detection and removal. **a** The main body is detected from the bee blob. **b** The main body is removed from the bee image (Color figure online)

Fig. 6 The result of pollen detection. **a** The pollen and other dust are detected. **b** The pollen and non-pollen blobs (Color figure online)

3.2 Features of Pollen Blobs

In the previous section, 273 pollen blobs are detected, but 3982 non-pollen blobs are also detected. Other pollen sac features can be used to eliminate many non-pollen blobs. The detected pollen blobs have two features that are different from non-pollen blobs.

The first feature is the area of the pollen region, which is the number of pixels inside the blob. However, a bigger bee blob causes the pollen blob to be bigger, often because the bee flies nearer the camera. To compensate for this, the area ratio (R_A) between the pollen blob and the bee main body (ellipse) blob is used as the feature:

$$R_A = \frac{A_p}{A_m} \tag{38}$$

A_p is the pollen blob area, and the A_m is the single bee main body blob area.

The second feature is the location of the pollen blob relative to the bee body. The pollen sac is beside the bee's body. The positions of the pollen sacs should be to the left and/or the right side of the bee's main body. The ends of the minor axis of the main body ellipse represent the left and right side of the main body.

Figure 7 shows the main body ellipse. It is assumed the head of the bee points in the direction of the positive x-axis. The orientation of the ellipse ($\theta \in (-90^\circ < \theta < 90^\circ)$) is the angle between the major axis of the ellipse and the positive x-axis. It is assumed $\theta < 0$ in the situation of Fig. 7. The point L_p is on the left side of the bee, which is at the top end of the minor axis of the ellipse in the figure; the point R_p is on the right side of the bee, which is at the bottom end of the minor axis of the ellipse in the figure. In Fig. 7, the point positions are:

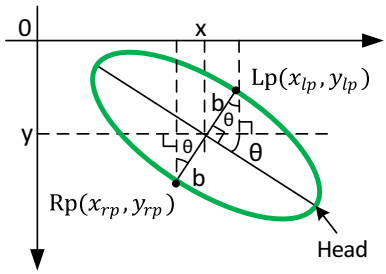

Fig. 7 The main body ellipse showing the left and right points of the minor axis (Color figure online)

$$Rp(x_{rp}, y_{rp}) = (x + b\sin\theta, y + b\cos\theta) \tag{39}$$

$$Lp(x_{lp}, y_{lp}) = (x - b\sin\theta, y - b\cos\theta) \tag{40}$$

where **x** and **y** are the centre position of the main body and b is the minor semi-axis length. The pollen sac blobs should be near the left point or right point. The distances between the left/right points and the pollen blobs' centre positions are calculated as:

$$D_{pr} = \sqrt{|Rp - P_b|^2} = \sqrt{(x_{rp} - p_{bx})^2 + (y_{rp} - p_{by})^2} \tag{41}$$

$$D_{pl} = \sqrt{|Lp - P_b|^2} = \sqrt{(x_{lp} - p_{bx})^2 + (y_{lp} - p_{by})^2} \tag{42}$$

where D_{pr} and D_{pl} are the distances between the detected blob and the *Rp* and *Lp* points. p_{bx} and p_{by} are the x and y coordinates of the blob position respectively. However, the distances vary with the size of the bee's main body, so they are calculated relative to the length of minor axis of the ellipse (q_{Rp} and q_{Lp}).

$$q_{Rp} = \frac{D_{pr}}{2b} \tag{43}$$

$$q_{Lp} = \frac{D_{pl}}{2b} \tag{44}$$

The minimum value of q_{Rp} and q_{Lp} ($Q_p = \min\{q_{Rp}, q_{Lp}\}$) is the second feature that is used for pollen detection. It is called the pollen position distance ratio.

These two features ("Position distance" and "Area ratio") are calculated for all of the 4255 blobs (273 pollen and 3928 non-pollen blobs) in order to find the optimal thresholds which can be used to discriminate pollen and non-pollen blobs. The results a plotted on Fig. 8, where each point represents a different blob. The x-axis is the "Position distance" (d-axis) and the y-axis is the "Area ratio" (a-axis).

From Fig. 8, pollen blobs are mainly distributed in the range 0.0–0.4 for "Position distance" and 0.01–0.08 for "Area ratio". The non-pollen blobs are concentrated in the range of 0.00–0.02 for "Area ratio" and 0.0–1.5 for "Position distance". While some of the pollen and non-pollen blobs are mixed, most of the non-pollen blobs can be separated from the pollen blobs. Two straight lines have

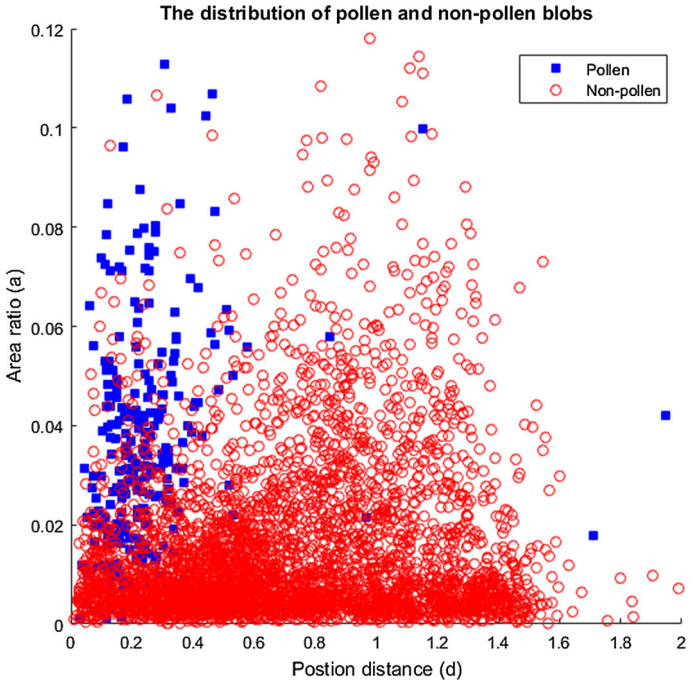

Fig. 8 The distribution of pollen and non-pollen blobs with position distance and area ratio (Color figure online)

been used to separate the pollen and non-pollen blobs on Fig. 8. Both lines cross the d-axis at point d_c. These two lines can be defined by the point d_c and the reciprocals of line slopes.

$$\mathrm{L_1} = \mathrm{a_n} - \frac{\mathrm{d_n} - \mathrm{d_c}}{\mathrm{m_1}} (\mathrm{m_1} > 0, \mathrm{d_n} \geq \mathrm{d_c}) \tag{45}$$

$$\mathrm{L_2} = \mathrm{a_n} - \frac{\mathrm{d_n} - \mathrm{d_c}}{\mathrm{m_2}} (\mathrm{m_2} < 0, \mathrm{d_n} < \mathrm{d_c}) \tag{46}$$

In Eqs. (45) and (46), L_1 and L_2 represent the two straight lines where $1/m_1$ and $1/m_2$ are the slopes of the two lines. The pollen blobs are concentrated in the region where $L_1 \geq 0$ and $L_2 \geq 0$ in the plane of the 2-D distribution (Fig. 8). This is the "two lines method" that is used to discriminate pollen and non-pollen blobs. By inspection, the ranges of possible values for the three line parameters are: $\mathrm{d_c} \in [0, 0.4]$, $\mathrm{m_1} \in [0.01, 11]$ and $\mathrm{m_2} \in [-20, -0.01]$. These three line parameters have optimal threshold values which can distinguish pollen and non-pollen blobs most effectively.

3.3 ROC for Discrimination

The optimal thresholds of the three line parameters are found by using the receiver operating characteristic (ROC) [16]. Using threshold values, the four possible outcomes for each blob are:

- True positive (TP): system detecting pollen for a pollen blob.
- False positive (FP): system detecting pollen for a non-pollen blob.
- False negative (FN): system detecting non-pollen for a pollen blob.
- True negative (TN): system detecting non-pollen for a non-pollen blob.

According to [16], the ROC uses sensitivity (SNS) and specificity (SPC) to test the thresholds of the line parameters. The sensitivity is defined as:

$$SNS = \frac{TP}{TP + FN} \tag{47}$$

The specificity is:

$$SPC = \frac{TN}{FP + TN} \tag{48}$$

The ROC graph parametrically plots SNS versus 1-SPC with the threshold as the varying parameter. The graph plots 1-SPC on the x-axis and SNS on the y-axis, to create a 2D space as in Fig. 9. This is the ROC space. On this space, changing a threshold from minimum to maximum will produce a curve from the left bottom corner to the top right corner. This is the ROC curve. If the analysis has a single parameter, there is one curve on the space. An example is shown in Fig. 9. For a single parameter, this curve is obtained by calculating SNS and SPC for 11 threshold values, to produce 11 points from (0,0) to (1,1).

Multiple parameters produce multiple curves on this 2D space, because finding the optimum combination of thresholds requires varying one parameter value for a number of combinations of values for the other parameters.

An ideal threshold would yield a point in the upper left corner at coordinate (0, 1) of the ROC space. This point corresponds to a perfect classification. However, none of the thresholds can reach this point in the situation of this research. Therefore, the optimal thresholds correspond to the point on ROC curve that is closest to the perfect classification point.

From Sect. 3.2, three line parameters need to be analysed. For each of the three line parameters, 51 thresholds were chosen, ranging from minimum to maximum. The ROC curves obtained by using the two lines method are shown in Fig. 10. There are 2601 (51 * 51) curves in the ROC space corresponding to the different combinations of values of m_1 and m_2. Each curve includes 51 ROC points corresponding to the 51 different values of d_c. In total, there are 132,651 (51^3) points on the ROC space. One point has the shortest distance to the perfect classifier (0, 1).

The optimum point is at (0.1185, 0.9121) with the shortest distance from (0, 1) being 0.1476. The y and x coordinates of the point correspond to a sensitivity of

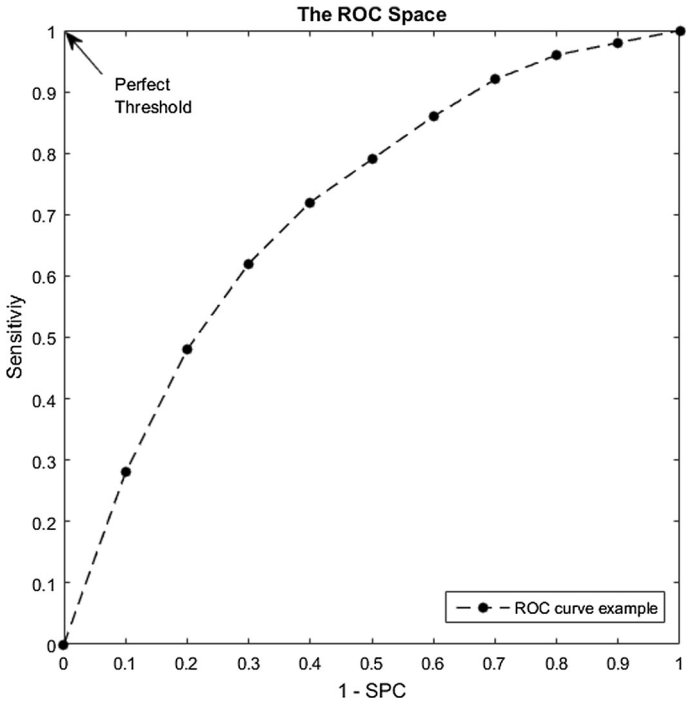

Fig. 9 An example of ROC space and curve

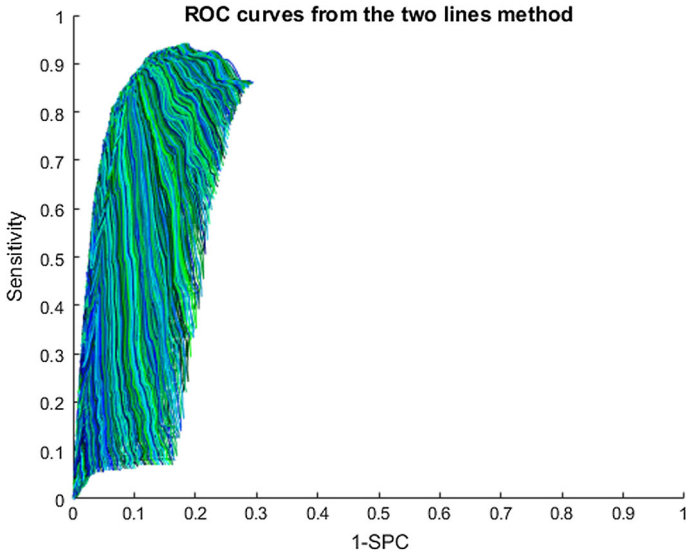

Fig. 10 The ROC curves of the two lines method (Color figure online)

0.9121 and a specificity of $1 - 0.1185 = 0.8815$. The optimal thresholds for the three line parameters are listed in Table 1.

From Table 1, the two optimal lines are:

$$L_1 = a_n - \frac{d_n - 0.208}{6.8238}(d_n \geq 0.208) \tag{49}$$

$$L_2 = a_n - \frac{d_n - 0.208}{-16.002}(d_n < 0.208) \tag{50}$$

The two lines are drawn on the (a–d) distribution map in Fig. 11 with blue and green colours respectively. The two lines cross on the d-axis at 0.208. The positive part is shaded in the figure. It can be seen that the positive part covers most of pollen blobs and the negative part includes most of the non-pollen blobs.

Table 2 displays the four outcomes of ROC. The true positive (TP) is 249, and the false positive (FP) is 472. In total, the two lines method identifies 721 blobs as

Table 1 Result of the two lines optimisation

m_1	m_2	d_c
6.8238	− 16.002	0.208

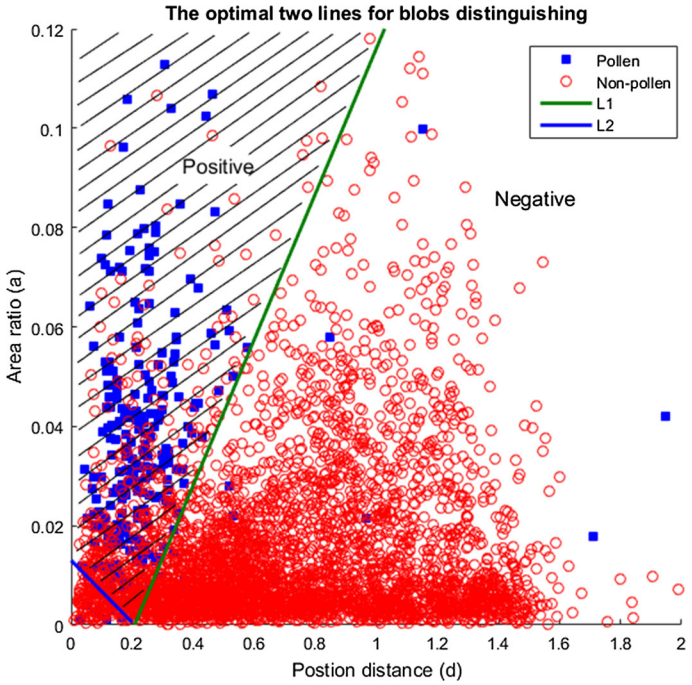

Fig. 11 The optimal two lines for discrimination of pollen from non-pollen blobs (Color figure online)

Table 2 ROC outcomes from the two lines method

TP	FN	FP	TN
249	24	472	3510

pollen. Of these blobs, 249 are actual pollen blobs, the other 472 are non-pollen blobs.

The two lines method significantly decreases number of non-pollen blobs. However, the pollen discrimination is still not very accurate because the FP of 472 is about twice of the number of the TP of 249.

The aim of this research is to count the pollen sacs on the video. All of these pollen and non-pollen blobs have been detected from individual bee images. Many images are of the same bees, recorded in different video frames. In the video, bees with pollen should be detected with pollen many times on the video frame flow, and bees without pollen should have a lower number of pollen detections. The next section will combine the bee tracking model and pollen detection model for pollen measurement.

4 Pollen Measurement Model

On the video, the purpose of the pollen measurement is to count the number of bees that are carrying pollen. Figure 12 shows the model of pollen measurement (or counting) on the video. As a bee is tracked on video, the model records the number of frames in which the bee individually appears on the video. This is the single tracking count. Bee appearances on the boundary of the video are ignored because the model does not allow analysis of these cropped images.

The bee tracking can be used to avoid multiple counting of pollen sacs, because the pollen sacs are detected on the individual bee images on each frame of the video. During the tracking, the model collects individual bee images and detects the pollen they are carrying. As the pollen detection model cannot find the pollen in the merged bee situation, these blobs cropped from the video are not analysed. In each frame of the video, once a bee's individual image is detected including pollen, the model increases the count of pollen detection for the bee. Because the bee is tracked on the video, the count can be accumulated with the detection of pollen on each frame. This is the bee's pollen count. When the bee disappears from the video, the model calculates the ratio of the pollen count to the single tracking count. This ratio indicates the probability of the bee having pollen.

However, if an individual bee fully appears on only one or two frames the ratio is not meaningful. Therefore, pollen is not counted on a bee appearing in less than three frames.

This ratio can be used to decide whether the bee has pollen and greatly improves the pollen count on the video. The threshold for the ratio is chosen from the ROC curve. In the first 500 frames of the sample video, 16 bees were carrying pollen sacs

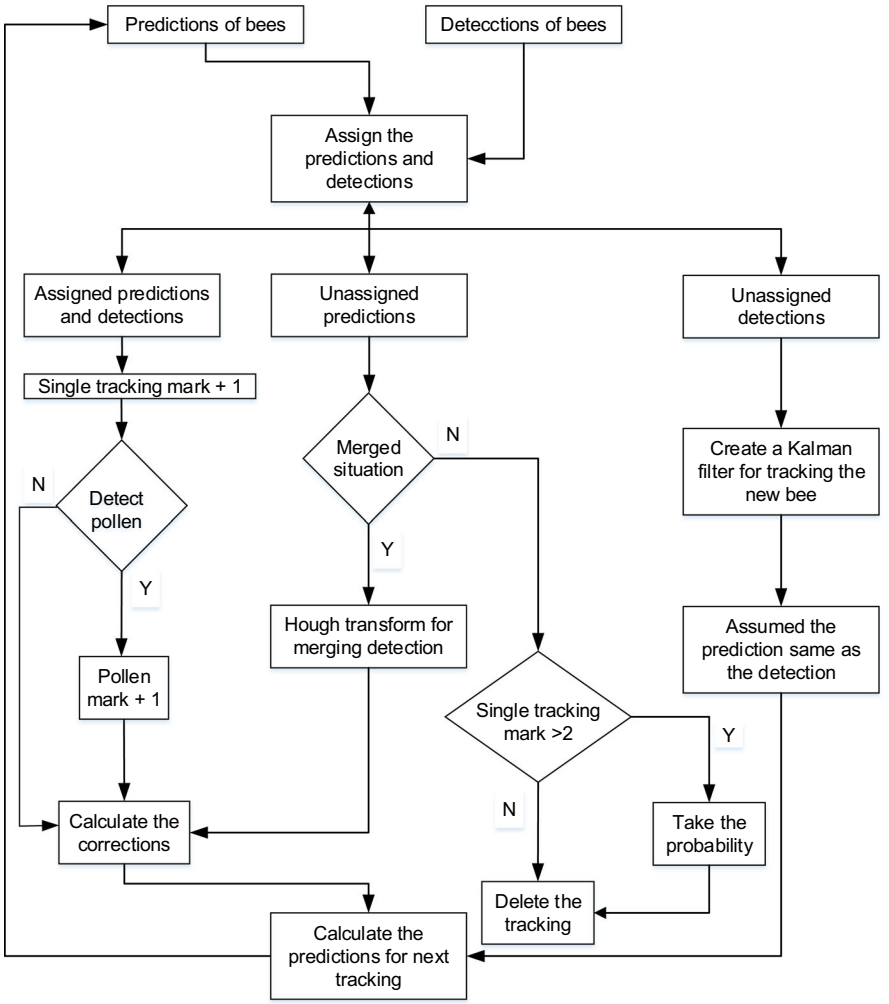

Fig. 12 The model of pollen measurement

and 85 bees did not have pollen. All of these bees appear individually on more than two frames. The ratio for each of these bees is used to choose 21 thresholds from 0 (0%) to 1 (100%). The ROC curve is shown in Fig. 13.

In Fig. 13, the point in the ROC space which is closest to the perfect classifier (0, 1) has a 1–SPC value of 0.059 and a sensitivity value of 0.875. The distance between the optimum point (0.059, 0.875) and the perfect classifier (0, 1) is 0.1381. The optimal ratio threshold is 0.65. Therefore, a bee that has a ratio greater than 0.65 is counted as carrying full pollen sacs. The results for the 16 pollen-carrying bees and 85 non-pollen bees are analysed using this threshold for the pollen counting result. This is shown in Table 3.

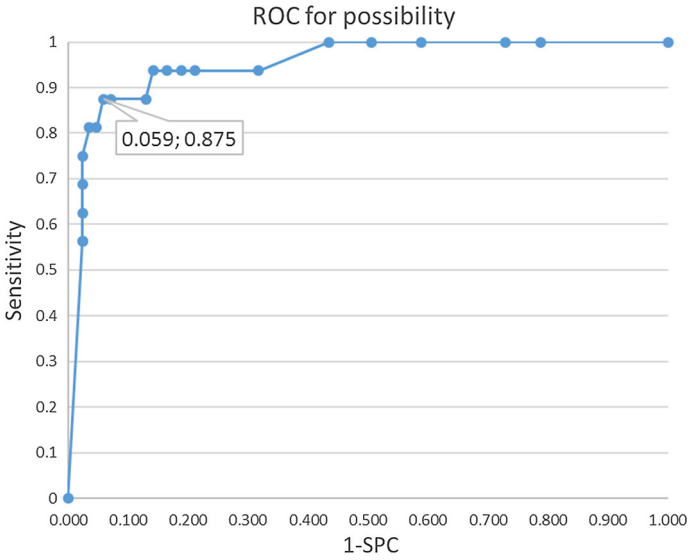

Fig. 13 The ROC curve for the pollen possibility (Color figure online)

Table 3 ROC result for the optimal threshold

Sensitivity	0.875	Specificity	0.941
TP	FN	FP	TN
14	2	5	80

From Table 3, the model calculates that 19 bees (14 + 5) were carrying pollen sacs during the first 10 s (500 frames) of the video. The actual number of pollen carrying bees was 16. The pollen count result is close to the actual count.

In summary, when a bee appears on the video, the pollen counting model checks whether the bee is carrying pollen on each frame until the bee disappears from the video. If the bee crosses the video boundary and disappears from the video, the model calculates the pollen-carrying ratio. If the ratio is greater than 0.65 (65%), the bee is identified as carrying full pollen sacs. The model counts the number of bees carrying pollen on the video with this process. Using this process, more video frames are used to test the pollen-counting model. This is described in the next chapter.

5 The Experimental Results

5.1 Results from Images

An example of a bee detection result is shown in Fig. 14. This is a result from one frame from a test video. The original video image is shown in Fig. 2 Sect. 1.

(a)

(b)

Fig. 14 The bee detection result. **a** Binary image showing bee blobs. **b** The bee detection and tracking result shown on the colour image (Color figure online)

Figure 14a shows the binary image which comes from the output of the bee detection. The white blobs are the bee bodies that are detected. This binary image is analysed to obtain the blob positions and size (height and width). The sizes of the blobs are used to draw bounding boxes around the bees on the colour image. These boxes show the bee detection result on the colour image of Fig. 14b which are the yellow bounding boxes. On Fig. 14b, seven bees are detected, and none of the shadows are detected. The bee detection is successful.

Figure 14b also displays the result of individual bee tracking. Bee tracking is indicated by adding indexes (bee IDs) to each bounding box and curves marking each bee's trajectory with different colours in the video. The red and green bounding boxes show the Kalman filter prediction and correction respectively. A tracking boundary is set up just in front of the hive entrance, which is shown by the purple line, so that the system only tracks bees in front of the green background area. This is because the area below the boundary line has complex colour information that affects the bee detection.

Figure 15 displays the merged bee tracking result. Figure 15a–h, as well as i–p, are taken from consecutive frames. Figure 15a is the situation before the two bees merged. Figure 15i shows the two bee blobs. These two bee blobs are used to obtain the bees' shapes for the Hough transform to use for the merged bee tracking. Figure 15b–g show the two bees merged. In Fig. 15j–o, the blob is one merged blob rather than two bee blobs in each frame. The merged blob on each frame is used to draw the blue bounding box on the corresponding colour image. The Hough transform is used on the merged blob to detect each bee's position. On the images of Fig. 15b–g, the green curves show the result of the Hough transform. These curves are the individual bee's shapes before their merging. This is a successful tracking example. The green curves (shapes) cover each actual bee shape well, and the curves change direction following the bees. In Fig. 15h, it can be seen that the two bees emerge with the same indexes they had in image Fig. 15a. The bees' flying trajectories follow the corresponding bees clearly. The tracking is successful even though these two bees were merged for six frames.

Pollen detection is applied to each individual bee image except for bees on the boundary and merged bees. Usually, a bee appearing on the boundary does not

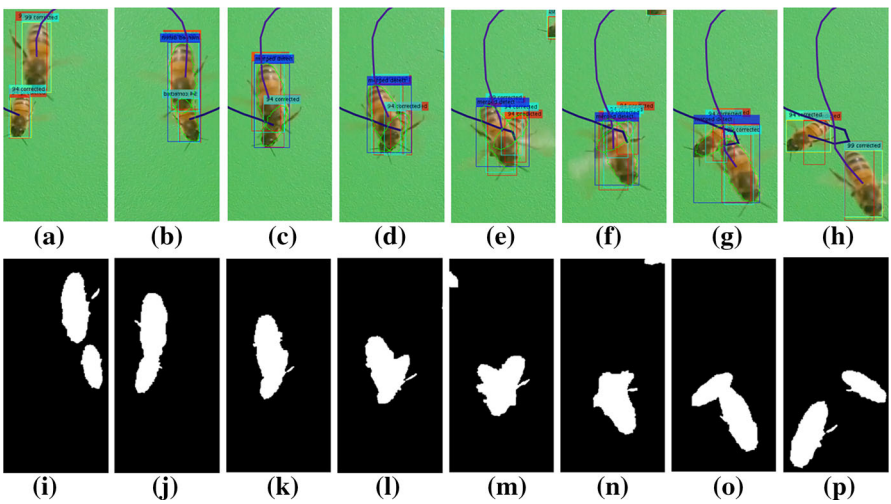

Fig. 15 A merged bee tracking example. **a–h** The colour images and tracking result on consecutive frames. **i–p** The blobs of merging bees (Color figure online)

appear as a full body. This is not suitable for the pollen detection model. Merged bee images are too complex to be used for pollen detection.

For example, in Fig. 14, seven bees are detected. Two of them appear on the boundary. Pollen detection will not be applied to these two bees. The other five bees are cropped out from the colour image of Fig. 2 by using the yellow bounding boxes on Fig. 14b. These five individual bee images are shown in Fig. 16a. Their corresponding blobs are cropped from the binary image of Fig. 14a. These blobs are displayed in Fig. 16b. Figure 16c shows the main body segmentation. It can be seen that the elliptical main bodies are removed from the images. The background is removed using the blobs of Fig. 16b. On Fig. 16c, the remaining colour parts

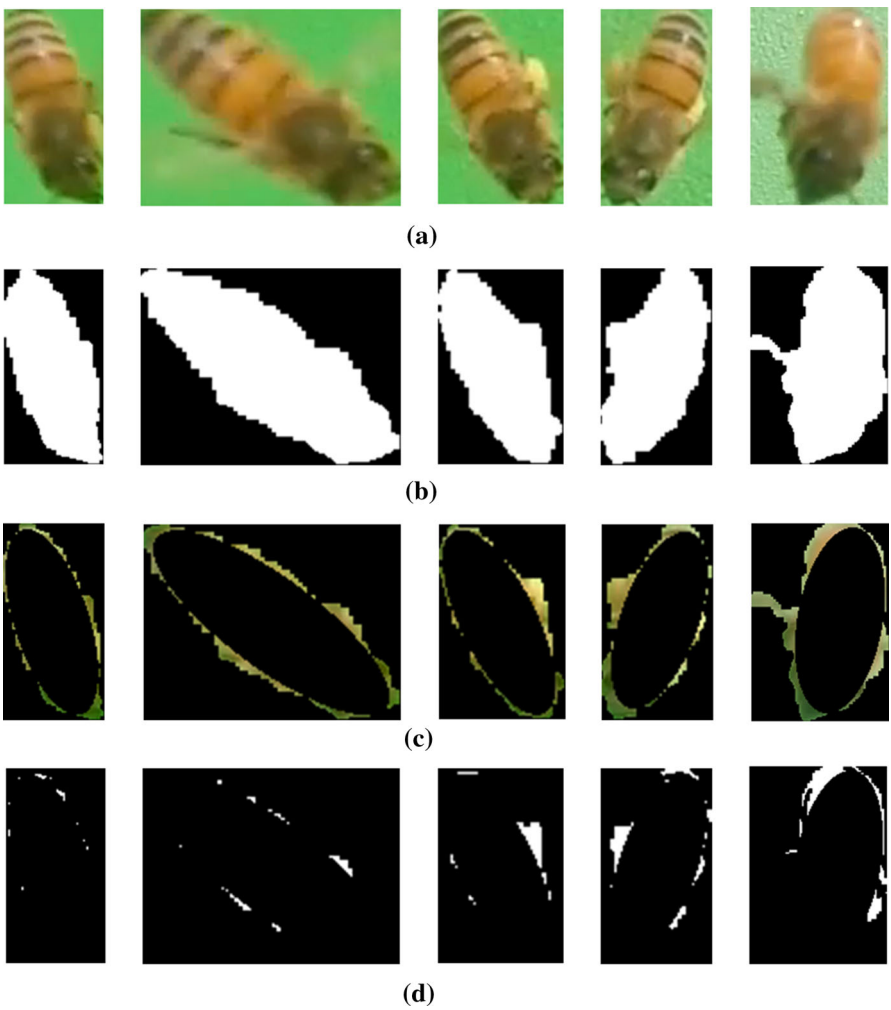

Fig. 16 An example of the pollen detection result. The five bees are from Fig. 2 in Sect. 1. **a** The 5 individual bee images. **b** The bee blobs from the binary image of Fig. 14a. **c** The main body and background segmentation. **d** The pollen detection result on the images (Color figure online)

include pollen sacs, legs, tail parts and head parts. These images are analysed using colour thresholding to find orange and white colours. The final pollen detection result is shown in Fig. 16d. All of the blobs on the images are candidate pollen blobs. Most of the blobs are small noise blobs. Two images clearly show pollen blobs. The right image of Fig. 16d shows a big blob which is the bee's tail part.

After pollen detection on the images, the ROC and two lines method (Sect. 3.3) is utilized to analyse each image of Fig. 16d to determine whether an individual bee image has pollen or not. On a video, the number of frames showing a particular bee with pollen are counted. In addition, the number of frames of this bee showing its full body is also counted. These two numbers are recorded by the system model (Fig. 12 in Sect. 4) until the bee disappears from the video. During the recording, the two numbers are displayed beside the bee on the video. These two numbers for each bee are displayed in Fig. 14b. The purple mark on the top of the bounding boxes shows the pollen situation of the bee. The two numbers are displayed on the purple mark in form of 'number1/number2'. The number1 shows the number of frames in which the bee shows pollen, and the number 2 displays the number of frames in which the bee is found flying individually. The pollen probability ratio is calculated as $\frac{\text{number1}}{\text{number2}}$. If this ratio is greater than 65% when the bee disappears from the video, the model adds one to the Pollen Count that is on the bottom left of the video.

5.2 Results from the Videos

On the monitoring videos, only a small number of bees carry pollen back to the beehive at the same time. It was difficult to get videos recording bees carrying pollen on the same day. Seven videos were chosen for the experiment. They are recorded from different times on different days. The recording environment is sunshine, so that the brightness and colour is similar among the videos. Each of the videos contain 3000 frames (1 min) for this experiment. The pollen measurement results are shown in Table 4, using the 65% threshold calculated from the ROC curve.

In Table 4, all the videos have very high specificity from 95 to 99%, because most bees are not carrying pollen on the video. In contrast, the sensitivity shows

Table 4 Pollen measurement on the videos

Video	Time	Frames	TP	FN	FP	TN	Sensitivity	Specificity
1	10:00	3000	29	9	10	408	0.76	0.98
2	10:46	3000	26	18	21	389	0.59	0.95
3	11:00	3000	55	37	19	930	0.60	0.98
4	11:26	3000	14	11	3	408	0.56	0.99
5	13:00	3000	24	29	6	533	0.45	0.99
6	14:00	3000	16	22	2	469	0.42	0.99
7	16:00	3000	9	39	1	1065	0.19	0.99

large variations for the different videos. Video 1 produced the best result with a sensitivity of 76%, but video 7 had only 19% sensitivity. The other five videos have sensitivities from 40 to 60%. The morning videos are generally better than the afternoon videos. Sensitivity is typically about 60% in the morning but around 40% in the afternoon. The main reason for the false negative results (FN) is that the pollen sacs are obscured by the bees' bodies for most of the time that the bees are recorded on the video.

In addition, Table 5 shows the measured number (TP + FP) and the actual number (TP + FN) of bees carrying pollen in these videos. The values are calculated from Table 4. It can be seen that the measured numbers are close to the actual numbers in videos 1–4. Particularly, videos 1 and 2 have near perfect measurement values. They are 39 and 47 respectively, and the actual numbers are 38 and 44. In videos 5 and 6, the measurement values are about half of the actual number. Video 7 has the most unsuccessful measurement, due to the high false negative (FN) value in Table 4.

In Sect. 3.3, the two lines method produced the results in Table 2. The pollen number measurement is 721, which comes from TP + FP = 249 + 472. However, the actual pollen number is 273, equal to TP + FN = 249 + 24. Because there is great number of false positives, the measured number has a large error. However, this result came from the counting of individual bee images. The pollen measurement on the videos is different because it includes bee tracking.

Comparing Table 5 and the result of the two lines method, the pollen measurement model including bee tracking greatly decreases the number of false positives. In addition, the sum of true positive (TP) and false negative (FN) is 338 over all of the videos. This is the actual number of bees carrying pollen. The sum of true positive (TP) and false positive (FP) is 235 over all the videos. This is the measured number of bees carrying pollen. The measured number is about 70% of the actual number.

Table 5 Measured and actual numbers

Video	Time	TP + FN	TP + FP
1	10:00	38	39
2	10:46	44	47
3	11:00	92	74
4	11:26	25	17
5	13:00	53	30
6	14:00	38	18
7	16:00	48	10
Total		338	235

6 Conclusion

This paper introduces a model to count the number of bees carrying pollen in monitoring video recordings. The model use motion detection and colour features to detect bees. Then it combines the Kalman filter and Hough transform to track bees. After that, the main bodies (head, abdomen and tail) of bees are removed on the cropped individual bee images. Possible pollen sacs are detected using colour thresholding, and the pollen sac blobs are discriminated from non-pollen sac blobs using the blob area and blob distance from the side of the bee. Finally, for each individual bee, the results obtained from all the bee's tracked images are used to decide whether the bee is carrying pollen or not.

The experimental results show the sensitivity is a maximum of 76%, and is mostly in the range of 40 to 60%. The pollen measurement model including tracking greatly decreases the number of false positives. Although the resulting measurement is not very accurate, it can provide beekeepers with an approximate measure of the pollen storage situation in the beehive.

In the pollen measurement from video, false negatives (FN) are another problem. The main problem causing false negatives (FN) is the pollen sacs being obscured behind the bee bodies. Therefore, it may be possible to improve pollen detection by using 3D video with two cameras recording video at different angles. Using two synchronized cameras recording videos from two different angles, rather than recording video with a vertical view, may identify pollen sacs more clearly than the video with a vertical view used in this research.

Acknowledgements The paper is supported by Darren Bainbridge (http://www.myapiary.co.nz, http://findmyhive.com), who provided access to beehives.

References

1. Desai, U. B., Merchant, S. N., Zaveri, M., Ajishna, G., Purohit, M., & Phanish, H. S. (2005). Small object detection and tracking: algorithm, analysis and application. In S. Pal, S. Bandyopadhyay & S. Biswas (Eds.), *Pattern recognition and machine intelligence* (Vol. 3776, pp. 108–117, Lecture Notes in Computer Science): Springer Berlin.
2. Wei, J. (2013). Small moving object detection from infra-red sequences. *International Journal of Image and Graphics, 13*(03), 1350014. https://doi.org/10.1142/S0219467813500149.
3. Baojun, Q., Tao, W., Bin, D., & Hangen, H. (2011). Fast detection of small infrared objects in maritime scenes using local minimum patterns. In *2011 18th IEEE international conference on image processing (ICIP), September 11–14, 2011* (pp. 3553–3556). https://doi.org/10.1109/icip.2011.6116483.
4. Vard, A., Jamshidi, K., & Movahhedinia, N. (2012). Small object detection in cluttered image using a correlation based active contour model. *Pattern Recognition Letters, 33*(5), 543–553. https://doi.org/10.1016/j.patrec.2011.11.012.
5. Qi, B., Wu, T., He, H., & Hu, T. (2011). Real-time detection of small surface objects using weather effects. In R. Kimmel, R. Klette & A. Sugimoto (Eds.), *Computer vision—ACCV 2010* (Vol. 6494, pp. 27–38, Lecture Notes in Computer Science): Springer Berlin.
6. Hsieh, F.-Y., Han, C.-C., Wu, N.-S., Chuang, T. C., & Fan, K.-C. (2006). A novel approach to the detection of small objects with low contrast. *Signal Processing, 86*(1), 71–83. https://doi.org/10.1016/j.sigpro.2005.03.020.
7. Campbell, J., Mummert, L., & Sukthankar, R. (2008). Video monitoring of honey bee colonies at the hive entrance. *Visual Observation & Analysis of Animal & Insect Behavior, ICPR, 8,* 1–4.

8. Kimura, T., Ohashi, M., Crailsheim, K., Schmickl, T., Odaka, R., & Ikeno, H. (2012). Tracking of multiple honey bees on a flat surface. In *2012 Fifth international conference on emerging trends in engineering and technology* (*ICETET*) (pp. 36–39). IEEE.
9. Chiron, G., Gomez-Krämer, P., & Ménard, M. (2013). Outdoor 3d acquisition system for small and fast targets. Application to honeybee monitoring at the beehive entrance. In *Workshop proceedings of GEODIFF 2013, in conjuction with VISIGRAPP 2013* (pp. 10–19).
10. Chiron, G., Gomez-Krämer, P., Ménard, M., & Requier, F. (2013). 3D tracking of honeybees enhanced by environmental context. In A. Petrosino (Ed.), *Image analysis and processing—ICIAP 2013* (Vol. 8156, pp. 702–711, Lecture Notes in Computer Science): Springer Berlin.
11. Wang, K., Liang, Y., Xing, X., & Zhang, R. (2015). Target detection algorithm based on gaussian mixture background subtraction model. In Z. Deng & H. Li (Eds.), *Proceedings of the 2015 Chinese intelligent automation conference* (Vol. 336, pp. 439–447, Lecture Notes in Electrical Engineering): Springer Berlin.
12. Cardani, D. (2001). *Adventures in hsv space*. Mexico: Laboratorio de Robótica, Instituto Tecnológico Autónomo de México.
13. Cheng, Y., & Collins, J. (2015). A model for honey bee tracking on 2D video. In *2015 International conference on image and vision computing New Zealand* (*IVCNZ*)*, November 23–24, 2015* (pp. 1–6). https://doi.org/10.1109/ivcnz.2015.7761542.
14. Nandashri, D., & Smitha, P. (2015). An efficient tracking of multi object visual motion using Hungarian method. *International Journal of Engineering Research and Technology, 4*(4), 1307–1310.
15. Chiu, S.-H., & Liaw, J.-J. (2005). An effective voting method for circle detection. *Pattern Recognition Letters, 26*(2), 121–133.
16. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874.