Bumblebees Detection and Tracking

Bernardo Miranda, Joaquín Salas,* and Pablo Vera

Abstract

Given the contradictory recent reports on whether there is a decline of insect pollinators, there is a clear need to develop more sophisticated monitoring systems in order to assess the quantity and variety of pollinators in a given environment. In this work, we explore an approach to the stability/plasticity dilemma to construct bumblebees tracklets, where the detector provides a stable, learned off-line, model of the object being sought, while an adaptable tracker keeps a record of both the object and the background to fill up the gaps in case of a detector miss. Through our experiments, we showed how the performance of the detector alone was enhanced with the addition of tracking.

1 Introduction

Shortly after the fall of 2006, there were worrisome reports[1] of a rapid and widespread decline of managed honey bee (*Apis mellifera L.*) colonies. This event was nicknamed Colony Collapse Disorder (CCD) [2] and its main trait was a rapid loss of adult worker bees. The fact is that human's food supply depends heavily on a few insect pollinators and a handful of plants. This is particularly troubling because within this fragile scenario there are currently frequent reports about declining populations of bumblebees[3] as well as observed migration of pathogens between pollinators[4].

Clearly, there is a need to develop monitoring systems to help assess the quantity and variety of pollinators in a given environment. Currently, this task is mostly performed using either direct observation [5] or offline video monitoring[6, 7]. Recently, [8] proposed a bumblebee detection algorithm based on the use of a Viola-Jones classifier[9]. Although the use of the algorithm represented a step forward in terms of the implementation of automatic analysis tools, this document presents a scheme where the combined use of detection and tracking improves the overall system performance.

In the next section we survey the literature on the topics of insect detection and tracking, with special emphasis on the work that has been done on computer vision strategies to monitor bees in general. Then, in §3 we explore how one can approach the plasticity-stability dilemma with the use of a classifier to initiate trackers and a discriminative tracker to fill the gaps between detection misses. Afterwards, in §4 we further detail how the detector and tracker are combined. Then, in §5 we show some experimental results to characterize the detector-tracker combination performance. Finally, we conclude the paper by discussing our results and plotting some possible directions of research.

2 Review of the Literature

Given its importance, the area of insect detection has been a very active research field. Some methods are especially suitable for the cases where the insects do not move, including the one developed by Xiao *et al.* [10] who presented a method to analyze butterflies. They make use of spectral regression to reduce the high dimension of the space used for classification. At their end, Huang *et al.* [11] proposed a method that extracts SURF descriptors of insect images. Similarly, Gao *et al.* [12] presented a method for the identification of insect species. They propose using Hu's invariants to extract the features of dragonfly wings.

In our case, we are interested on analyzing flying bumblebees while they are functioning as pollinators, as they approach flowers. In this context, some strategies applied for determining insect behavior include the following. In [13], Kahn *et al.* use Eigentracking, *i.e.*, Principal Component Analysis, to track bees using a particle filter[14]. Along these lines, Tsukamoto *et al.*[15] proposed a particle filter for a variable-appearance object that allowed to successfully

*The authors are with the Instituto Politecnico Nacional, Cerro Blanco 141, Colinas del Cimatario, Queretaro, Mexico, 76090. Joaquin Salas, `jsalasr@ipn.mx`, is the corresponding author. Thanks to Paul Riley for his comments to the manuscript. This research was partially supported by SIP-IPN through grant number 20121642.

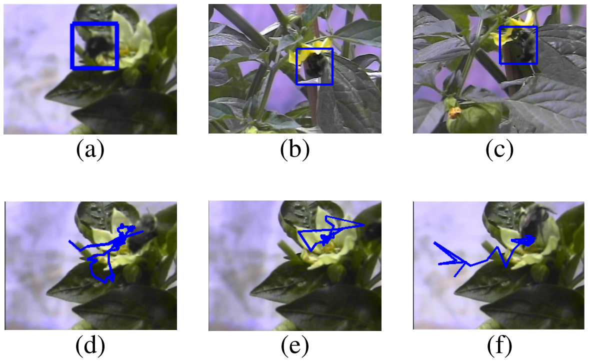

Figure 1. Bumblebee detection and tracking. In (a)-(c), we show three examples of bumblebee detection. Then in (d)-(f) we illustrate the tracklet described by the bumblebee. Our method to combine detection and tracking methods improves the performance over the detection alone.

track blinking fireflies in the dark. Specifically related to bees, Veeraraghavan *et al.* [16] proposed a model to track dancing bees in a hive.

In this research, we observed bumblebees and flowers as they are placed in a large cage. Our long term objective is to characterize the interaction between pollinators and flowers, such that the actions of the former leads the latter to become fruit. To that end, in [8], we developed a Viola-Jones classifier [9] to detect bumblebees. However, the use of detection alone results in fragmented tracklets [17], *i.e.*, a chronological set of observations, that makes difficult the analysis of trajectories. In [18] the problem of tracking is surveyed and a prime problem is identified as the construction of adaptable models to avoid drifting. In this work, we explore an approach to the stability/plasticity dilemma[19] to construct bumblebees tracklets. Thus, while the detector provides a stable, learned off-line, model of the object being sought, an adaptable tracker keeps a record of both the object and the background to fill up the gaps in case of a detector miss. Although perhaps more sophisticated methods could be applied, we study the use of the Gu and Tomasi[20] tracker because of its high performance, fast execution, and available code.

3 Detection and Tracking Components

Here, we describe the previously reported experience with the use of a Viola-Jones classifier for bumblebee detection and also review the Gu-Tomasi tracker.

The popularity of the Viola-Jones classifier [9] is based on its speed, simplicity, effectiveness, and, more recently, the availability of efficient implementations.

Algorithm 1 Combining tracking and detection

Input: A steady flow of images $\{I_k\}$
Output: A set of tracklets $\mathscr{T}$
$\mathscr{T} \leftarrow \emptyset$ {*algorithm's output*}
$\mathscr{S} \leftarrow \emptyset$ {*internal tracklets register*}
$k \leftarrow 0$ {*current frame index* }
loop
  Update tracklets in $\mathscr{S}$ using the Gu-Tomasi tracker {*The state of the tracklet is* **track**}
  Mark every $\tau \in \mathscr{S}$ as unvisited
  $D \leftarrow$ Detect bumblebees in I_k using the Viola-Jones classifier
  if $D \neq \emptyset$ **then**
    for $d \in D$ **do**
      if $d \bigcap (\tau \in \mathscr{S}) \neq \emptyset$ **then**
        mark τ as visited {*The state of the tracklet is* **collapse**}
      else
        create a tracklet τ and add it to $\mathscr{S}$ {*The state of the tracklet is* **birth**}
      end if
    end for
  end if
  for unvisited $\tau \in \mathscr{S}$ **do**
    if τ has been tracked for more than ς frames without a detection **then**
      $\mathscr{S} \leftarrow \mathscr{S} - t$ {*Remove tracklet from the internal list. The state of the tracklet is* **dead**}
      $\mathscr{T} \leftarrow \mathscr{T} \bigcup t$ {*Add tracklet to the output list*}
    end if
  end for
  $k \leftarrow k+1$ {*Process the next image*}
end loop

In [8] a bumblebee detection based on the Viola-Jones classifier algorithm was presented. In that work, a series of classifiers was constructed and tested that show that with an 18 stage classifier it is possible to obtain a *false positive rate* of less than 0.01% while showing a *true positive rate* of around 85%. Given this performance, the resulting tracklets are fragmented. Thus, the aim of a tracker will be to profit from the temporal coherence provided by the image sequence.

In [20], Gu and Tomasi define a tracking algorithm that represents each image with a set of features, updates a bag of features that is used to represent the object of interest using nearest-neighbor classification [21], and looks for its position in the next frame using Efficient Subwindow Search [22]. The method uses SIFT descriptors [23] to create an appearance model. Each image is represented with a set of Scale-invariant Feature Transform (SIFT) key points $V(I) = \{(\mathbf{x}_i, \mathbf{v}_i)\}$, for $i = 1, \ldots, n$. Here $\mathbf{x}_i \in \mathbb{R}^2$ is the set of positions, and $\mathbf{v}_i \in \mathbb{R}^d$ is the set of SIFT features. Once a detection occurs, a window W is defined. To represent the set of key point descriptors of I within the window W, it is defined as $\Theta(W; I)$. Considering $\mathscr{B} \subset \mathbb{R}^d$ as the background model, Gu and Tomasi propose to update the object model $\mathscr{O}_k \subset \mathbb{R}^d$ using the previous object model $\mathscr{O}_{k-1}$ and the current window W_k as:

$$\mathscr{O}_k \leftarrow \mathscr{O}_{k-1} \cup F_\lambda[\Theta(W_k; I_k), \mathscr{O}_{k-1}, \mathscr{B}], \tag{1}$$

where F_λ enriches the features set of the current model incorporating good ones and filtering out bad ones.

4 Combining Detection and Tracking

The tracker developed by Gu and Tomasi[20] has the ability to adapt well to fast changing object appearance while the Viola-Jones classifier[9] has the potential to provide support in case of drifting. Therefore, we combine the two methods into a single unified framework that results in a highly flexible and adaptable method for detection and tracking.

Consider a sequence of images $\{I_k\}$. Let us assume that at frame I_j, m bumblebees are detected by the Viola-Jones classifier inside the window D_j^i, for $i = 1, \ldots, m$. For each detection D_j^i, a Gu-Tomasi tracking process, in the form of tracklet τ_s, is started. As the sequence advances, the Gu-Tomasi tracker estimates new locations for the object of interest $T_{j+l}^{i'}$, for $l = 0, \ldots, \varsigma$. Note that at the first detection corresponds to the first position of the tracklet, *i.e.*, $D_j^i = T_j^{i'}$. In our case ς is the maximum number of frames that a bumblebee can be tracked without a detection. Whenever a detection does not occur for the last ς observations, the tracker gives up, these ς observations in the tracklet τ_s are discarded, and the tracklet τ_s is added to the tracklets set $\mathscr{T}$. Thus, each tracklet $\tau_i = \{O_i^1, \ldots, O_i^r\}$ has observations that could come from either the Viola-Jones classifier D_j^i or the Gu-Tomasi tracker $T_j^{i'}$. The tracklet can be in any of the following four states. (1) **birth:** A tracklet τ_s is born when its first detection D_k^i occurs; (2) **track:** This happens when a detection does not occur and the position of the tracklet endpoint $T_{j+l}^{i'}$ is defined by the Gu-Tomasi tracker; (3) **collapse:** Both detection D_j^i and tracking $T_j^{i'}$ occur for the same bumblebee; and (4) **death:** This occurs when after ς frames there have been no detections for the tracklet. The algorithm 1 details the above description.

5 Results and Discussion

To compare the relative performance of detection plus tracking versus the detection only, we implemented the algorithm described. As described in [8] and using their database of images, a set of Viola-Jones classifiers, with different number of stages, were trained using 1,237 positive samples and 1,000 negative samples. Afterwards, the different classifiers were tested with a sequence of 4,000 images containing five bumblebees visiting a flower. The sequence contained 1,668 positive samples and 2,332 negative samples. Among the different classifiers, the 18 stages classifier was chosen because it gave a high positive rate (0.85) and a low false positive rate (0.004). Some figures illustrating the detection of bumblebees using this classifier are shown in Fig. 1(a)-(c). The same sequence of 4,000 frames later used to test the detection plus tracking strategy and compares it with the detection-only algorithm. There are two details to consider in our implementation. On one hand, the tracking process gives up whenever the detector does not return a positive result in five consecutive frames. Therefore, the last five frames of the tracklet are dropped and do not accumulate false positives. Also, whenever the tracklet consisted of only the initial detection (presumably a false positive) and five consecutive frames without detection, the whole tracklet was dropped without accumulating in the count of false positives. A summary of the results is provided in the Table 1. We illustrate the tracklets generated with several examples obtained from the sequence of images used for experimentation in Fig. 1 (d)-(f). The code was programmed using MATLAB mex files and Microsoft Visual C++. We used an Intel core $i3$ machine, with each core running at 2.13 GHz. The average processing time for the detection was 0.01 s/frame, while the tracking stage added an additional 0.05 s/frame.

Conclusion

In this paper, we describe an effective strategy to add temporal constraints via tracking to a bumblebee detector. From our experiments, we have shown how the performance of the detector alone was enhanced with the addition of tracking. Furthermore, the additional information provided via tracking could prove to be most helpful in tasks such as behavior assessment or in the determination of the relationship between flowers and insect pollinators. As an example, given the position of the flowers, which could possibly be estimated with a method such as the one described in [24], we could determine the quality of the pollination process and the health of a bumblebee colony by measuring how often a flower is visited and how long a bumblebee stays in the flower.

With a True Positive Rate of 0.96 and a False Positive Rate of 0.001, our results seem very promising. In the

Table 1. A comparison between the elements of the confusion matrix in the decision process to detect bumblebees.

	Detection	Detection plus Tracking
True Positives	1,426	1,617
False Negatives	253	62
False Positives	6	1
True Positive Rate	0.85	0.96
False Positive Rate	0.004	0.001

near future, we plan to continue our experiments with longer image sequences and a more diverse variety of flowers with the objective of assessing more precisely the performance of our detection plus tracking strategy. Additionally, we plan to develop systems to study the interaction between insect pollinators and flowers, thus extending the capabilities of current monitoring methods.

References

[1] F. D. Cox, F. Maryann, and H. J. Ostiguy Nancy, "Falldwindle Disease: Investigations into the Causes of Sudden and Alarming Colony Losses Experience by Beekeepers in the Fall of 2006," *International Journal on the Biology of Stress*, 2006.

[2] J. Evans, C. Saegerman, C. Mullin, E. Haubruge, B. Nguyen, M. Frazier, J. Frazier, D. Cox-Foster, Y. Chen, and R. Underwood, "Colony Collapse Disorder: A Descriptive Study," *PLoS One*, vol. 4, no. 8, p. e6481, 2009.

[3] S. Cameron, J. Lozier, J. Strange, J. Koch, N. Cordes, L. Solter, and T. Griswold, "Patterns of Widespread Decline in North American Bumblebees," *Proceedings of the National Academy of Sciences*, vol. 108, no. 2, p. 662, 2011.

[4] A. Core, C. Runckel, J. Ivers, C. Quock, T. Siapno, S. DeNault, B. Brown, J. DeRisi, C. Smith, and J. Hafernik, "A New Threat to Honey Bees, the Parasitic Phorid Fly Apocephalus borealis," *PloS one*, vol. 7, no. 1, p. e29639, 2012.

[5] B. Lingren and D. Lenker, "Insect Trap and Assembly for Capturing and Monitoring Insects," 2003, US Patent 6,516,558.

[6] C. Lortie, A. Budden, and A. Reid, "From Birds to Bees: Applying Video Observation Techniques to Invertebrate Pollinators," *J. Pollination Ecology*, vol. 6, 2012.

[7] R. Steen, A. Aase, and A. Lene, "Portable Digital Video Surveillance System for Monitoring Flower-visiting Bumblebees," *J. Pollination Ecology*, vol. 5, no. 13, pp. 90–94, 2011.

[8] B. Miranda, J. Salas, and A. Torres, "Automatic Visual Detection of Bumblebees with a Viola-Jones Classifier," *Journal of Pollinators Ecology-Submitted*, 2012.

[9] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *CVPR*, vol. 1, 2001.

[10] L. Xiao-Lin, H. Shi-Guo, Z. Ming-Quan, and G. Guo-Hua, "KNN-Spectral Regression LDA for Insect Recognition," in *International Conference on Information Science and Engineering*, 2009, pp. 1315–1318.

[11] S. Huang, X. Li, M. Zhou, and G. Geng, "SURF-Based Multi-scale Resolution Histogram for Insect Recognition," in *International Conference on Artificial Intelligence and Computational Intelligence*. IEEE, 2009, pp. 445–448.

[12] Y. Gao, H. Song, X. Tian, and Y. Chen, "Identification Algorithm of Winged Insects based on Hybrid Moment Invariants," in *International Conference on Bioinformatics and Biomedical Engineering*, 2007, pp. 531–534.

[13] Z. Khan, T. Balch, and F. Dellaert, "A Rao-blackwellized Particle Filter for Eigentracking," in *CVPR*, vol. 2, 2004, pp. II–980.

[14] M. Isard and A. Blake, "Condensation-conditional Density Propagation for Visual Tracking," *IJCV*, vol. 29, no. 1, pp. 5–28, 1998.

[15] Y. Tsukamoto, Y. Matsumoto, and T. Wada, "Tracking a Firefly-a stable Likelihood Estimation for Variable Appearance Object Tracking," in *ICPR*, 2008, pp. 1–4.

[16] A. Veeraraghavan, R. Chellappa, and M. Srinivasan, "Shape and behavior encoded Tracking of Bee Dances," *T-PAMI*, pp. 463–476, 2008.

[17] J. Prokaj, M. Duchaineau, and G. Medioni, "Inferring Tracklets for Multi-object Tracking," in *CVPR*, 2011, pp. 37–44.

[18] S. Salti, A. Cavallaro, and L. Di Stefano, "Adaptive appearance modeling for video tracking: Survey and evaluation," *T-IP*, vol. 21, no. 10, pp. 4334–4348, 2012.

[19] S. Grossberg, "Competitive learning: From interactive activation to adaptive resonance," *Cognitive Science*, vol. 11, no. 1, pp. 23–63, 1987.

[20] S. Gu, Y. Zheng, and C. Tomasi, "Efficient Visual Object Tracking with Online Nearest Neighbor Classifier," in *ACCV*, 2010.

[21] O. Boiman, E. Shechtman, and M. Irani, "In Defense of Nearest-neighbor based Image Classification," in *CVPR*, 2008, pp. 1–8.

[22] C. Lampert, M. Blaschko, and T. Hofmann, "Efficient Subwindow Search: A Branch and Bound Framework for Object Localization," *T-PAMI*, vol. 31, no. 12, pp. 2129–2142, 2009.

[23] D. Lowe, "Object Recognition from Local Scale-invariant Features," in *ICCV*, 1999, p. 1150.

[24] P. Vera, J. Salas, R. Manduchi, and E. Sanchez, "Deteccion automatica de la apertura y el cierre de las flores de Mammillaria mathildae Kraehenb. & Krainz," *Journal of the Professional Association for Cactus Development*, pp. 89–97, 2010.