

Tracking Bees - A 3D, outdoor small object environment

Author

Estivill-Castro, V, Lattin, D, Suraweera, F, Vithanage, V

Published

2003

Conference Title

2003 INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, VOL 3, PROCEEDINGS

DOI

[10.1109/ICIP.2003.1247421](https://doi.org/10.1109/ICIP.2003.1247421)

Rights statement

© 2003 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Downloaded from

<http://hdl.handle.net/10072/1792>

Griffith Research Online

<https://research-repository.griffith.edu.au>

TRACKING BEES - A 3D, OUTDOOR SMALL OBJECT ENVIRONMENT

V. Estivill-Castro D. Lattin F. Suraweera

V. Vithanage

School of Computing and Information Technology
Griffith University
Brisbane 4111 QLD Australia

Long Pocket Laboratories
CSIRO - Plant Industry
Brisbane 4068 QLD Australia

ABSTRACT

The automatic tracking of bees while pollinating macadamia trees is important for the understanding of reproductive biology and fruit yield. We present techniques for tracking these small targets in an uncontrolled illumination environment and where the background is not fixed. Our results indicate we can track, based on color, bees moving in 3D in paths over 3s long (72 frames).

Keywords: Segmentation, motion detection and estimation, image sequence processing, application of image processing technology to plant biology.

1. INTRODUCTION

Applications of image segmentation, motion detection in a sequence of frames and object tracking have focused on tracking relative large objects that travel in restricted fashion. A typical example is tracking vehicles from another vehicle [1]. These systems have achieved real-time performance; but large number of pixels in a blob correspond to moving objects. Thus, methods can afford extracting features (lines and corners) to identify object flow [2]. Also, the objects do not change direction rapidly (cars are essentially restricted to travel in 2D space), so the clusters of directional vectors can be updated smoothly over several frames. Optical flow techniques are computationally expensive and regarded as unrealistic in many real-time situations unless specialized hardware is available [3].

We present algorithms for tracking bees (small insects) flying around macadamia trees (an outdoor environment with uncontrolled illumination). We use inexpensive equipment and segmentation on color. The background changes as wind and other factors move the flowers and leaves on the scene. Also, the camera itself is subject to vibration. This challenging application is our focus. Our research is part of a larger project where obtaining reliable data of insects acting as pollinators is crucial for optimization processes of agricultural products. So far, the only attempt (that we are aware of) to record automatically the behavior of insects involves tracking 2D paths from ants in an in-door, illumination controlled environment [4]. The tracking of such small

objects on rather complex moving backgrounds precludes an approach that would use edge detection algorithms to identify regions and then analyzes differences between regions rather than pixels in nearby frames.

2. THE APPLICATION

Over 30% of humans' total diet worldwide derives directly or indirectly from insect pollination of plants. All horticultural crops and some field crops respond to pollination. Reproductive factors other than pollen compatibility are involved in pollination success (such factors include effectiveness of bees in enhancing pollen flow under orchard conditions). Also, integrating these factors can result in recommendations to maximize yields by improved management and/or orchard design for macadamia. However, the understanding of all these factors and their influence is limited.

In an orchard situation, the behavior of the pollinating agents, the bees, is dependent on the flower opening pattern and the duration of their attractiveness. When the flowers are visible and producing nectar and pollen, the bees need to visit them. This is the period corresponding to pollen receipt on the stigma and for which we collect our data.

Individual flowers are attractive for 3 days [5]. Within a day, the flower opening peaks in the early afternoon [5] which can correspond to one of the peaks of the bimodal insect activity [6]. The European honeybee and the native bee, *Trigona* sp., are the most important pollinators of macadamia [6, 5]. The honeybee is effective over long distances while the *Trigona* is effective in promoting near neighbor pollination [7]. While the nectar attracts the bees to macadamia, it is the availability of pollen that determines the actual pollen flow [8].

2.1. Experimental Setting

The number and movement of honeybees during flowering are typically investigated through manual data collection method and random sampling at random times. This approach is time consuming and resource intensive. Our goal is to develop a more automated approach. The video stream

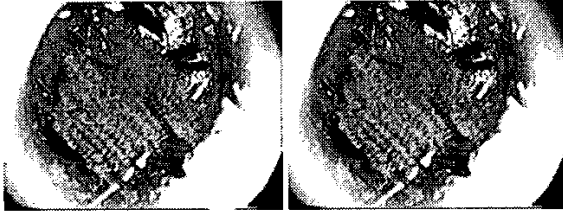


Fig. 1. Two consecutive frames.

was captured over a wireless link and a camera placed in a macadamia orchard. The image shifts position due to wind blowing around the flowers even if the camera is fixed. Wind shakes the camera as well. The speed and the direction of bees is difficult to predict. Bees can move quickly across the image sequence and also hover motionless across image sequences. Noise is introduced from the camera and transmission of the video across wireless medium. The image quality is not high. In fact, the environment requires the use of inexpensive color CCD cameras. The camera is exposed to severe weather conditions and thus, it was boxed into a waterproof container. The placement in an orchard eliminated fixed power sources and thus boxed batteries had to be provided. The batteries were large to support long recording times. The camera had to be fixed in a position that reduced its movement as much as possible but oriented toward a cluster of flowers. The camera has a focal length of approximately 30cm and a field of vision of approximately 160° . The wireless link covered approximately 100m. The receiver allows the videocassette reorder to store intervals of over one minute every 10 minutes. The video-capture device allowed images of 320×240 pixel resolution at 24 frames per second. The capture of the video was stored digitally as three channel RGB video file in '.avi' format. Figure 1 shows two consecutive frames where one bee is flying over clear background.

3. THE TECHNIQUES

One of the main lessons from years of research in computer vision is that "segmentation of an image into 'meaningful' parts is almost always the first step in any image analysis process" [9]. We are interested in detection and tracking of moving bees in image sequences. This falls within the area of Change Detection [10](CD). The task involves finding significant differences between the frames in input video stream under the varying illumination and camera movement conditions. Little research has attempted this in outdoor, time-varying conditions [10] although object identification based on color is probably one of the most studied problems in computer vision and pattern recognition. It is

believed that for controlled illumination environments (like in-door situations), object recognition is solved by classical statistical methods like maximum-likelihood classifiers that model color variations as Gaussian noise. For un-controlled illumination scenarios (like outdoors) we face added complexity. The apparent color of an object is affected by the illuminant color, the illumination geometry and the viewing angle [11]. Also, shades, cloud cover and other factors [11] have stimulated the use of other features and more costly sensors (like infra-red, radar, etc). Nevertheless, object identification by color continues to be of interest in outdoor settings [12]; specially due to the low cost of the required image-capture equipment.

Our techniques are based on CD methods known as frame differencing but will be based on color. Thus, we explain frame differencing methods successful in tracking cars [13], traffic monitoring [3] and object tracking in general [14, 15]. The techniques assume that the camera is fixed and the image is not moving from frame to frame [16, 17, 18]. Because we do not have a constant background, we use two-frame differencing. That is, we use $N \times M$ digitized images were the pixel for the t -th frame $I_t(x, y)$ and its predecessor $I_{t-1}(x, y)$ produce a binary image. The *simple difference* (SD) method [19] assumes images are gray scale values (for example, $I_t(x, y) \in \{0, \dots, 255\}$) and computes the absolute difference $D_t = |I_t(x, y) - I_{t-1}(x, y)|$. Then, a threshold filter of value $T \in \{0, \dots, 255\}$ is applied: $SD_t(x, y) \begin{cases} 0 & \text{if } D_t(x, y) < T \\ 1 & \text{otherwise.} \end{cases}$ Changes to illumination and noise affect the quality of the results of this method. These problems are reduced with the *deviate model* (DM) method [20] (also known as geo-differencing). Here, local statistics are computed by considering (square $n \times n$) regions R_t and R_{t-1} centered at (x, y) for the respective input images [21]. We have RGB values for each pixel $I_t(x, y)$; thus, we have $\{R_t(x, y), G_t(x, y), B_t(x, y)\}$. The approach proposed here combines four aspects.

- 1.- The use of a color projection as the leading feature to track a moving object.
- 2.- The locality of information around the object (spatial locality).
- 3.- The historical construction of a background image by weighted average (an evolving statistical model [22]).
- 4.- Domain knowledge on the size and aspect of the objects of interest.

3.1. Color projection

Consider SD for frame differencing or much more sophisticated methods (like binary statistical morphology [10]). Under the fixed camera assumption, a simple fact is that, if an object has moved between two frames, then something changes in the pixels corresponding to the object. Thus, we profit that we have 3 channels, and thus, we can trace simul-

taneously $R_t(x, y)$, $G_t(x, y)$, $B_t(x, y)$. Thus, our method monitors the three absolute differences $|R_t(x, y) - R_{t-1}(x, y)|$, $|G_t(x, y) - G_{t-1}(x, y)|$ and $|B_t(x, y) - B_{t-1}(x, y)|$ as well as a Minkowsky metric on this space of 3 features: $M_t(x, y) = ((|R_t(x, y) - R_{t-1}(x, y)|^p + |G_t(x, y) - G_{t-1}(x, y)|^p + |B_t(x, y) - B_{t-1}(x, y)|^p)^{1/p}$ (note that larger p corresponds to monitoring the largest change in a channel). Assume for now that the algorithm uses high differences in M as its starting directing feature. Then, the algorithm may tracing a group of 2 to 5 pixels that exhibit high difference on M because they correspond to a bee in motion and in clear view. However, when the bee flies behind a flower, values for M on the bee and its neighboring pixels become too similar. But, values in R (or perhaps the other 2 colors) show large discrepancy between the bee and the flower. This is because the flower offers some transparency as well as some shade on the bee. Because of our domain knowledge (described later), the algorithm switches its directive feature. From that moment, pixels with high discrepancy on R are highlighted. This switch of feature is triggered by loss of significance in discrepancy on the current leading feature, high discrepancy in another feature in very close proximity of the group of pixels highlighted in recent frames by the current feature. We have implemented this in our system with reasonable success using SD; note, however, that it could be extended to any frame differencing method.

3.2. Spatial locality

Again, consider SD as the starting point. Determining a threshold that would adjust to illumination conditions and applicable to all pixels in the image automatically seems truly difficult. However, we take advantage of the fact that our moving objects are very small. Thus, we can use a region centered on the recent small blob of highlighted pixels as the source of information. This locality is applied in two ways. First, rather than testing if a feature, say R has $|R_t(x, y) - R_{t-1}(x, y)| > T$ (where T is a global threshold), we identify if $|R_t(x, y) - R_{t-1}(x, y)|$ is significantly large with respect to the distribution of $|R_t(u, v) - R_{t-1}(u, v)|$ for pixels (u, v) in the local region R centered at (x, y) . Second, the region updates locally an evolving model described next.

3.3. Statistical Evolving Model

Collins et al [3] and Balch et al [4] use a Statistical Evolving Model (ESM) to automatically generate a fixed background. The approach involves a weighted average, where if at pixel (x, y) the current generated background is β_t then the new generated background is $\beta_{t+1} = \lambda\beta_t + (1 - \lambda)I_t(x, y)$ (where $\lambda = \tau \times f$ and f is the frame rate and τ regulates how quickly the generated background adapts to intensity changes). We not only apply this approach to create features

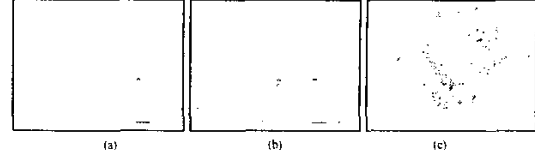


Fig. 2. SD works well on the clear.



Fig. 3. EM on each color channel reveals other bees.

that not only analyze change between consecutive frames, but also, change between current frame and background.

3.4. Domain knowledge

We incorporate domain knowledge in several aspects. The most immediate ones are large objects classified as having no interest. This usually correspond to entire flowers vibrating or swung by the wind. Similarly, very small objects (one pixel or 2) are removed as noise. Most importantly, we influence the analysis of the current frame and its predecessor by focusing on where are the tails of the current paths of bees that are being followed. In fact, this is an important component of our system and we extend it into the future. That is, we not only match paths from the past but also from the future. The simple principle we are applying here is that moving bees should be traceable objects if we play our video stream forward or if we play it backward.

4. RESULTS

The results of applying SD with $T = 105$ (Fig 2 (a)) and $T = 75$ (Fig. 2 (b)) detect well a bee flying over the clear background. However, bees over the flowers are not detected and lowering the value of T shows movement of leaves and not the other bees (Fig 2 (c)). By using ESM and the color channels we see that movement not detected on the blue channel is picked up by the green channel (for an example, see Figure 3). Thus, applying our techniques we obtain paths as shown in Figure 4. Here, the adaptive threshold T is computed over a region of 40 pixels around each pixel. The ESM is such that I_{t-3} contributes less than 1%. Blobs less than 3 pixels are removed.

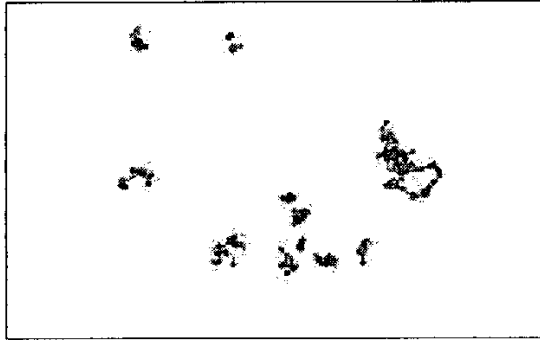


Fig. 4. Resulting paths on a segment of 120 frames.

5. FINAL REMARKS

Some problems still remain. A small stigma or a leave vibrating overtime is indistinguishable from a bee path except by the fact that remains fixed over the same set of pixels in the image. In Figure 4, some results (two bottom and two top) are not actually bees. Nevertheless, in the right most path, the bee flies over and behind flowers. It corresponds to a path of a bees moving in 3D for over 3s long (72 frames).

6. REFERENCES

- [1] S.M. Smith, "ASSET-2: Real-time motion segmentation and object tracking," *Real-Time Imaging*, vol. 4, 21-40, 1998.
- [2] D.J. Barron, J. L. Fleet and S.S. Beauchemin, "Performance of optical flow techniques," *Int. J. Computer Vision*, vol. 12, 42-77, 1994.
- [3] R. Collins, et al, "A system for video surveillance and monitoring: VSAM final report," 2000. Carnegie Mellon U. Pittsburgh, US.
- [4] T. Balch, Z. Khan, and M. Veloso, "Automatically tracking and analyzing the behavior of live insect colonies," *AGENTS-01*, Montreal, 2001, ACM Press.
- [5] T. A. Heard, "Pollinator requirements and flowering patterns of macadamia integrifolia," *Aust. J. Bot.*, vol. 41, 491-497, 1993.
- [6] V. Vithanage and D. A. Ironside, "The insect pollinators of macadamia and their relative importance," *J. Aust. Inst. of Agric. Science*, vol. 52, 155-160, 1986.
- [7] A. M. Montalvo and J. D. Ackerman, "Relative pollinator effectiveness and evolution of floral traits in *spathiphyllum friedrichsthalii* (araceae)," *American J. of Botany*, vol. 73, 1665-1676, 1986.
- [8] V. Vithanage and T. J. Douglas, "Honeybee pollination of macadamia: floral rewards and their effect on pollen flow," *J. Apicultural Research*, vol. 26, 261-269, 1987.
- [9] A. Rosenfeld, "From image analysis to computer vision: An annotated bibliography 1955-1979," *Computer Vision & Image Understanding*, vol. 84, 298-324, 2001.
- [10] G.L. Foresti, "Real-time detection of multiple moving objects in complex image sequences," *Int. J. Image Systems Technology*, vol. 10, 305-317, 1999.
- [11] B.K.P. Horn, *Robot Vision*, MIT Press, MA, 1986.
- [12] S.D. Buluswar and B.A. Draper, "Color recognition in outdoor images," *ICCV 6th Int. Conf. Computer Vision*. 1998, 171-177, IEEE Computer Society,
- [13] G.L. Foresti, V. Murino, and C. Regazzoni, "Vehicle recognition and tracking from road image sequences," *IEEE T. Vehicular Technology*, vol. 48, 301-317, 1999.
- [14] I. Haritaoglu, D. Harwood, and L. Davis, "Real-time surveillance of people and their activities," *IEEE T. Pattern Analysis and Machine Intelligence*, vol. 22, 809-830, 2000.
- [15] S. Young, M. Forshaw, and Hodgetts, "Change detection for perimeter surveillance," 1998. UCL Image Processing Group.
- [16] P. Rosin and T. Ellis, "Image difference threshold strategies and shadow detection," *British Machine Vision Conf.*, 1995, 347-356.
- [17] P. Rosin, "Thresholding for change detection," *British Machine Vision Conference*, 1997, 212-221.
- [18] T. Ellis and M. Xu, "Object detection and tracking in an open and dynamic world," *2nd Int. Workshop on PETS*, 2001.
- [19] R. Jain, D. Miltzer, and H.H. Nagel, "Separating non-stationary from stationary scene components in a sequence of real world tv-images," *Int. Conf. on Artificial Intelligence*, 1977, 612-618.
- [20] K. Skiestad and R. Jain, "Illumination independent change detection for real world image sequences," *Computer vision, graphics, and image processing*, vol. 46, 375-390, 1989.
- [21] J.C. Russ, *The image processing handbook*, CRC Press, Boca Raton, FL, 3rd edition, 1999.
- [22] R. Collins, A. Lipton, and T. Kanade, "A system for video surveillance and monitoring," 1999. Carnegie Mellon U. Pittsburgh, US.