Tracking Bees in a Hive

Ashok Veeraraghavan and Rama Chellappa
Dept. of ECE, University Of Maryland, College Park, MD 20742
vashok@cfar.umd.edu and rama@cfar.umd.edu

Behavorial research in the study of the organisational structure and communication forms in social insects like the ants and bees has garnered tremendous impetus in recent years [1][2]. Usually, when such an experiment to study these insects is setup, the insects in an observation hive are videotaped. The hours of videotape is then manually studied and hand-labelled. This task of manually tracking the insect motions in videos and then labelling the video data takes up the bulk of the time and effort of such experiments. In this paper, we provide methodologies for behavior based tracking of insects in such videos and show its application to the problem of tracking bees in a hive.

Most of the general approaches for tracking [3][4][5] do not directly adapt well to tracking insects because they exhibit very specific forms of motion. These videos contain significant occlusions. Moreover, there are also several other very similar bees around the bee being tracked. All these make the tracking task very challenging. [6] has proposed the use of a Rao-Blackwellized particle filter for tracking and has shown results on tracking bees. But this method cannot extract relevant physical parameters like the orientation of the abdomen with respect to the thorax etc. In this paper, we present an approach for automatic tracking and behavior analysis of bee movements.

A. *Shape and Behavior Model*

Bees like all other insects are made of three parts, -the head, the thorax and the abdomen. The shape model of the bee is made of three ellipses, one for each body part. The dimensions of the various ellipses are fixed during initialization. The location of the bee and its parts in any frame can be given by five parameters- namely, the location of the center of the thorax(2 parameters), the orientation of the head, the orientation of the thorax and the orientation of the abdomen. Tracking the bee over a video essentially amounts to estimating these five model parameters($\mathbf{X} = [x_1\ x_2\ x_3\ x_4\ x_5]'$) for each frame.

Bees exhibit a variety of behaviors[1] depending upon the nature of the work they are performing. We model the probability distributions of location parameters $\mathbf{X}$ for certain basic motions($m_1 - m_4$). In our implementation, we modelled four different motions- 1) Moving straight ahead, 2) Turning, 3) Waggle, and 4) Motionless. We use appropriate Gaussian pdfs (p_{m1}, p_{m3}, p_{m4}) for the straight, waggle and motionless while we use a mixture of two Gaussians for modeling turning motion.Each behavior is now modelled as a Markov process on these motions, i.e., each behavior is characterized by a prior probability for each of these motions and a transition probability matrix between these motions.The difference between motion models and behavioral models is the range of time scales at which modeling is done. Motion models typically model the probability distribution (pdf) of the position in the next frame as a function of the position in the current frame. Instead, behavioral models capture the probability distribution of position over time as a function of the behavior that the tracked object is exhibiting. We believe that the use of behavioral models presents a significant layer of abstraction that is able to capture the variety and complexity of the motions exhibited by the insects.

B. *Tracking Algorithm*

We address the tracking problem as a problem of estimating the state X_1^t given the observations Z_1^t. Since both the state transition model and the observation model are non-linear, filtering procedures like the Kalman filter are inadequate. This Bayesian estimation of the state given observations can be done recursively using particle filters. Particle filters provide a method of recursively updating the posterior pdf $P(X_t/Z_1^t)$ as a set of N weighted particles $\{X_t^{(i)}, \pi_t^{(i)}\}_{i=1}^N$.

The state transition and the observation models are given by,

$$\textit{State Transition Model: } X_t = F_B(X_{t-1}, N_t) \tag{1}$$

$$\textit{Observation Model: } Z_t = G(Y_t, X_t, W_t) \tag{2}$$

In our approach the state model is equivalent to our behavioral model. The observation model is a mixture of Gaussians. We have manually selected 5 frames from the video and extracted the appearance of the bee in those 5 frames. These serve as multiple color exemplars($A_1, , .., A_5$). The appearance of the bee in any given frame is assumed to be Gaussian centered around one of these five exemplars. Therefore, given the appearance of a particle ($Z_t^{(i)}$), the weight for the particle ($\pi_t^{(i)}$) is updated as

$$\pi_t^{(i)} = p(Z_t^{(i)}/X_t^{(i)}); \tag{3}$$

$$p(Z_t^{(i)}/X_t^{(i)}) = \min_{j=1,2,..,5} p(Z_t^{(i)}/X_t^{(i)}, A_i); \tag{4}$$

$$p(Z_t^{(i)}/X_t^{(i)}, A_i) = L(Z_t^{(i)}; A_i, \sigma^2) \tag{5}$$

where, $L(Z_t^{(i)}; A_i, \sigma^2)$ is a normal density with mean A_i and diagonal covariance matrix $\sigma^2 * I_{3dX3d}$ (3d since there are d pixels in the appearance model and the RGB components are treated independently).

Cateogary: Visual Processing and Pattern Recognition Preference: Oral

C. Results

We conducted tracking experiments on two video sequences of bees in a hive. In both video sequences there was a forager performing the waggle dance.We tracked the dancer in both videos. Similar results were obtained for both video sequences. In all our simulations we used 300 to 600 particles. Both sequences were over 250 frames long. We were able to track the dancer for the entire length of the video sequence without any missed tracks. Moreover, we were able to extract parameters like the orientation of the various body parts during each frame. We used these parameters to automatically identify the waggle portion of the dance. We also verified this estimate manually and found it to be robust and accurate. Automatic extraction of the orientation of the abdomen during the waggle dance is important because the waggle portion of the dance encodes information such as the distance and the direction of food source. Specifically, the orientation of the waggle axis indicates the direction of the food source.

Figure 1 shows the structural model of the tracked bee superimposed on the original image frame. The results are best viewed in color since the tracking algorithm had color images as observations. The figure shows the top five tracked particles(blue being the best particle and red being the fifth best particle). As is apparent from the sample frames the appearance of the dancer varies significantly within the video. There is also significant occlusions in some frames. Frames 41,170,172 and 187 show the ability of the tracker to maintain track during partial occlusions. In fact, in frame 172, occlusion forces the posterior pdf to become bimodal (another bee in close proximity). But we see that the track is regained when the bee emerges out of occlusion in frame 175. In frame 187, we see that the thorax and the head of the bee is occluded while the abdomen of the bee is seen. Therefore the estimate of the abdomen is very precise. Since the thorax is not seen we see that there is high variance in the estimate of the orientation of the thorax and the head. Structural modeling has ensured that inspite of occlusion, only physically realizable orientations of the thorax and the head are maintained.

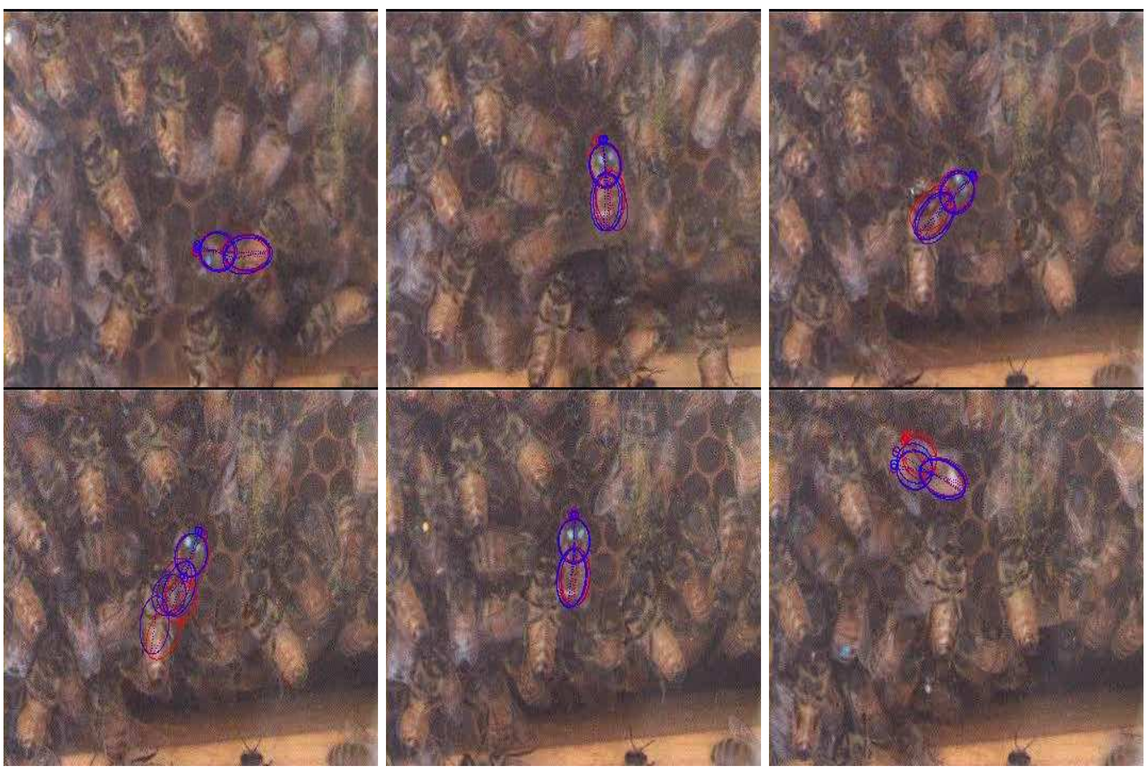

Fig. 1. Sample Frames from a tracked sequence of a bee in a beehive. (Shows the top five particles in each frame) Frame Numbers rowwise from the top left : 41,90,170,172,175 and 187

REFERENCES

[1] V. Frisch, *The Dance Language and orientation of bees*. Cambridge MA:Harvard University Press.

[2] M. Srinivasan, S. Zhang, M. Lehrer, and T. Collett, "Honeybee navigation en route to the goal: visual flight control and odometry," *The Journal of Experimental Biology*, vol. 199, pp. 237–244, 1996.

[3] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean-shift," *Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 142–149, 2000.

[4] M. Isard and A. Blake, "Contour tracking by stochastic propogation of conditional density," *European Conference on Computer Vision*, pp. 343–356, 1996.

[5] S. Zhou, R. Chellappa, and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters," *IEEE Trans. on Image Processing*, 2004.

[6] Z. Khan, T. Balch, and F. Dellaert, "A rao-blackwellized particle filter for eigen tracking," *Conference on Computer Vision and Pattern Recognition*, 2004.