Rigid Body Reconstruction for Motion Analysis of Giant Honey Bees using Stereo Vision

Michael Maurer, Matthias Rüther and Horst Bischof
Institute for Computer Graphics and Vision (ICG), Graz University of Technology
maurer/ruether/bischof@icg.tugraz.at

Gerald Kastberger
Institute for Zoology, Karl-Franzens University Graz
kastberger@uni-graz.at

Abstract

Zoologists are interested in the rapid movements of giant honeybees. Especially the mass-movement of bees during the defense behavior is of interest. State of the art approaches measure the motion of a single bee using a laser vibrometer, which does not provide geometric information on mass-motion like speed, intensity, starting point or mesh density. To overcome this problem, a vision based measurement system is proposed. A portable stereo setup using two high resolution cameras with high frame rates is designed to acquire image sequences of a defense wave in an outdoor environment. The functionality of the acquisition setup has been proven at an expedition to Nepal. Additionally, a framework to segment bees and reconstruct their motion is presented. The challenging task of solving the corresponding problem is solved using reduced graph cuts, to get accurate matches in the presence of repetitive patterns. An evaluation of the method has been done by comparison to manually labeled data.

1. Introduction

Giant honey bees (lat. apis dorsata) belong to the species of honey bees and live in the mainly forested areas of southern and southeastern Asia, like Nepal. The subspecies apis dorsata dorsata that was observed in this project has the second largest individuals among all honey bees. Typical workers are about 17 to $20\mathrm{mm}$ long. A colony of giant honey bees consists of up to $100,000$ individuals, living in an open-nest. The nest is made up of a single comb with up to two meters in horizontal span and about one meter in vertical span. The nests are usually built in exposed places far off the ground, like overhanging cliffs, trees or buildings (see Figure 1 and 2). Adult bees cover the comb in multiple layers, loosely fixed to the comb. In literature this construction is addressed as bee curtain, and is parted in two regions. One shows locomotion of individuals, while the other one remains stable. In this zone, which is of special interest to us, the bees mostly remain still and are uniformly oriented with the heads up.

Figure 1. Giant honey bee colony located at the backside of a hotel.

One method of defense against enemies is the defense waving, also called shimmering behavior. It is a mass event, comparable to the Mexican wave of people during sports events, and propagates over the bee curtain at a speed of about one meter per second. The waves are generated by coordinated rising of the abdomen of neighboring bees (see Figure 3) [1], [5].

From the viewpoint of zoologists a number of geometric parameters of the nest are of interest. These are:

- the mesh width, i.e. distance between neighboring

Figure 2. Giant honey bee colonies located at a 20m high water tower.

Figure 3. Bee cluster performing a wave by synchronized rising of their abdomen. The time between the images shown is 83ms.

bees on the curtain,

- the motion of the wave in 3D space,
- the speed in the different directions and
- the participation of second layer bees that are bees under the surface of the curtain.

Additionally, it is of interest to discriminate between active bees, which participate in the wave propagation, and passive bees, which are just dragged along by the others. Technically, to be able to solve these requirements, a segmentation of individual bees, a three-dimensional (3D) reconstruction of the bee curtain, and tracks of individual bees are desired. Having a 3D reconstruction of the bees over time, their movement can be evaluated. This evaluation also includes the participation of second layer bees (below the 'surface 'layer) that can only be observed by global movements of the bee curtain.

The technical challenge hereby lies in the field of view, defined by the dimensions of the curtain, the high wave propagation speed, the remote location, and the passive measurement principle. An active observation would possibly disturb the bees and force a counter-attack.

Stereo reconstruction was selected as a suitable method. An image acquisition system has been constructed which allows to monitor a complete nest from one side, at a synchronized frame-rate of 60fps. No extra illumination is required and the system is autonomously powered. From a methodological viewpoint, we faced the challenge of segmenting thousands of bees per image, establishing correct stereo correspondences on a very repetitive bee curtain, and tracking segmented individuals over several hundred frames.

2. Data Acquisition

We constructed a stereo camera setup with the following restrictions in mind:

- The bee cluster has a size of about $700 \times 700 \times 100\text{mm}^3$ (height $\times$ width $\times$ depth), that should be in the measurement range.
- Cameras cannot be mounted arbitrarily close to the cluster, to avoid unwanted interaction and counter-attacks. A clearance distance of two meters has to be maintained.
- The size of an individual is approximately $20\text{mm} \times 6\text{mm}$ (length $\times$ diameter). In an image the diameter should cover at least 10px.
- The reconstruction error in z-direction should not exceed one millimeter.
- An abdomen flip lasts for about 0.2s and should be resolved with 10 frames.
- The recorded sequence should contain one complete wave, so continuous acquisition time must be greater than one second.
- The stereo setup has to be placed outdoors, possibly in a high, remote location. It has to be portable, flexible and withstand influences like dust.
- The setup has to be self powered and should be able to operate for one working day.

A resolution of 4Mpx per camera was selected. According to the requirements, 700mm will fit to 1728px and one pixel will correspond to roughly 0.4mm in the real world. A bee with a diameter of 6mm will be represented by about 15px. The cameras are able to capture 60fps and so resolve an abdomen flip with 12 frames.

Two cameras at a frame rate of 60fps result in a data stream of $480\text{MB}/\text{s}$ that has to be stored. The captured image sequence is buffered to RAM and written to disc later. This leads to RAM constraining the capture time. A total of 8GB of RAM results in an acquisition time of 15s. The setup is powered by a 12V PC power supply and a car battery of 100Ah. So an operation time of seven hours can be reached. The resulting stereo setup is shown in Figure 4.

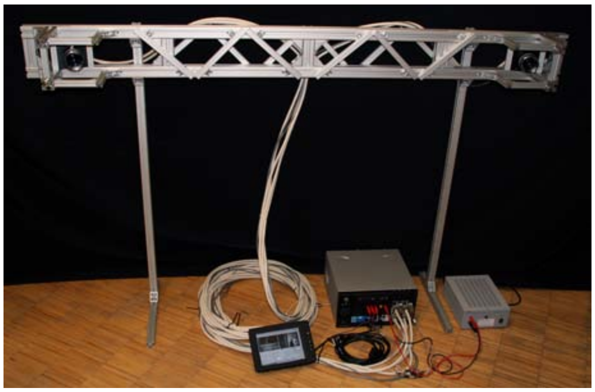


Figure 4. Image acquisition setup consisting of two cameras mounted to a camera stand, the industrial PC, a touch screen and an external power supply.

3. Methodology

To be able to analyze the defense waves, individual bees have to be segmented, matched, tracked and triangulated based on the stereo images.

Figure 5. Input image.

Having a close look at an input image (see Figure 5), the bees cover the comb in multiple layers and overlap closely. Heads and the thoraxes contrast weakly from the background and the semi-transparent wings cover parts of neighboring bees, which poses a challenge for segmentation.

An evaluation of various segmentation methods, including MSERs [2], shape prior segmentation [4] and template matching, led to the conclusion that template matching is the most efficient and accurate solution to the problem [3]. Segmentation of individuals was done by template matching using normalized cross correlation (NCC). To be robust to variations in size and orientation, a number of scaled and rotated templates were matched, using a winner-takes-all strategy. An exemplary result is shown in Figure 6.

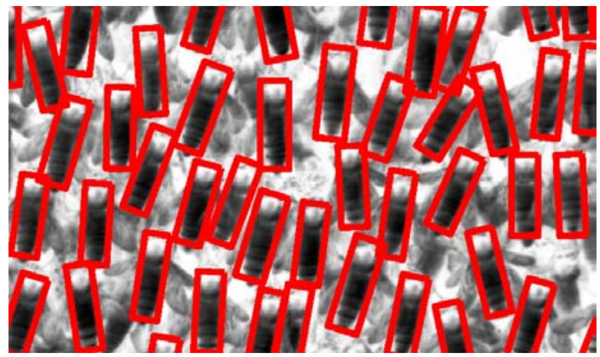


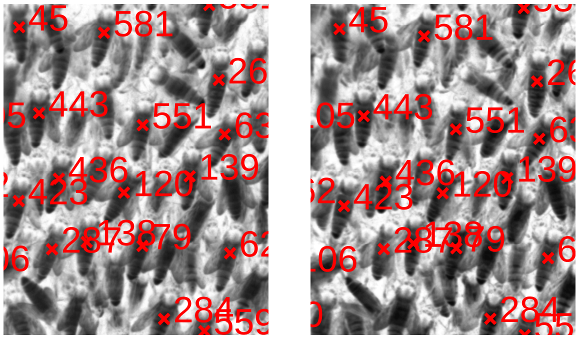
Figure 6. Segmentation of the single bees using template matching.

Establishing stereo correspondences between segmented individuals is still challenging, because of the apparently high similarity among them - otherwise segmentation by template matching would not have been successful.

Having a set of bees in discrete positions, and an initial estimate of disparity by robustly fitting a plane to the cluster, we formulate the matching process as a discrete optimization problem, and solve it by using reduced graph cuts, as proposed by Zureiki et al. [6]. Exemplary results are shown Figure 7.

In a final step, individuals should be tracked over time in the stereo images. Figure 8 illustrates the abdomen flip of a single bee and shows the apparent problems. These are the rapid flip of the abdomen, motion blur, and the extreme change in appearance. By tracking a bee at its thorax, we have to deal with occlusions caused by the raised abdomen.

Work on tracking is still in progress. We propose to use an adaptation of the Lukas-Kanade algorithm, where we seek to track the thorax position from the last n frames in every new frame. Again, the tracking result with highest confidence wins. In contrast to the naive approach of tracking only between consecutive frames, we are hereby more robust and exploit the symmetry in



(a) Camera 1 (b) Camera 2

Figure 7. Corresponding bees resulting from the minimum cut.

Figure 8. Sequence of a single bee flipping the abdomen at a frame rate of 60fps.

the abdomen motion. We further enforce the epipolar constraint by projecting corresponding tracks in the left and right image on the mean epipolar line (i.e. mean y-coordinate in rectified images).

A qualitative result is given in Figure 9, showing the z-motion of three individual bees over a time period which covers two consecutive defense waves.

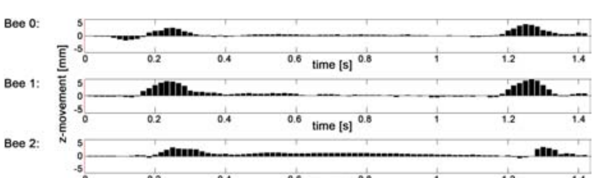


Figure 9. Z-movement of three single bees tracked over two consecutive defense waves.

4. Preliminary Results and Conclusion

In this paper a non-invasive vision system for monitoring a cluster of giant honey bees is presented. For the first time it is possible to analyse the mass behavior of these insects in-situ. The image acquisition system was successfully applied during an expedition to Nepal, where a total of 3.6TB of image data was acquired.

Using the proposed reconstruction approach, it is possible to robustly segment over 580 individuals in an area of $500 \times 500\text{mm}^2$. The correspondence problem under the presence of repetitive patterns is addressed using reduced graph cuts. There, using a stereo image pair with 630 and 521 segmented bees respectively, we retrieved 436 correspondences, 98% of which were correct by manual inspection.

The tracking problem is addressed using an adaptive template matching approach, but research is still in progress there. Having the positions of the single bees in 3D after triangulation, it is possible to perform an automated analysis of the z-movement and the mesh width in the first layer.

Future work will include an exhaustive processing and evaluation of the image data gathered in the last expedition, which sums up to 521 stereo movies. Processing time for a single movie currently lies at $14hrs$, which will also be improved in the near future.

References

[1] G. Kastberger. Communication and defense behavior of the asian giant honeybee apis dorsata. http://www.apis-dorsata.info. last visited: 2010.01.30.

[2] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *In British Machine Vision Conference*, volume 1, pages 384–393, 2002.

[3] M. Maurer. Rigid body reconstruction for motion analysis of giant honeybees using stereo vision. Master's thesis, Graz University of Technology, 2010.

[4] M. Werlberger, T. Pock, M. Unger, and H. Bischof. A variational model for interactive shape prior segmentation and real-time tracking. In *International Conference on Scale Space and Variational Methods in Computer Vision (SSVM)*, Voss, Norway, June 2009. to appear.

[5] Wikipedia. Apis dorsata. http://en.wikipedia.org/wiki/Giant_honey_bee. last visited: 2010.01.30.

[6] A. Zureiki, M. Devy, and R. Chatila. Stereo matching using reduced-graph cuts. In *ICIP07*, pages I: 237–240, 2007.