Image-Based Pose Estimation for Analyzing Cricket-Robot Interaction Behavior

Kuniaki Kawabata$^{1,2}$, Hitoshi Aonuma$^3$, Satoru Takahashi$^4$, Koh Hosoda$^5$ and Jianru Xue$^2$

$^1$RIKEN-XJTU Joint Research Unit, RIKEN
2-1 Hirosawa, Wako City, Saitama 351-0198, Japan
$^2$Institute of Artificial Intelligence and Robotics, Xi’an Jiaotong University
28, Xianning West Road, Xi’an, Shaanxi 710049, China
$^3$Research Institute for Electronic Science, Hokkaido University
N7W12, Kita-Ward, Sapporo City, Hokkaido 001-0020, Japan
$^4$Department of Intelligent Mechanical Systems Engineering, Kagawa University
2217-20, Hayashi-cho, Takamatsu City, Kagawa 761-0396, Japan
$^5$Graduate School of Information Science and Technology, Osaka University
1-3, Machikaneyama-cho, Toyonaka City, Osaka 560-8531, Japan
E-mail: kuniakik@riken.jp

Abstract This paper describes an image-based agent’s pose estimation method for supporting experiments on cricket-micro mobile robot interactions. Crickets modify their behavior selection tendency as a result of experience and the context based on the interaction with other crickets. Such a behavior selection mechanism with a tiny brain is a typical example to understand the adaptability of animals. Behavior observations in such interaction experiment sometime impose a burden on the experimenter. Therefore, in our previous works, we reported the development of a visual tracking method for two agents. However, it only tracked the positions of the agents. In this paper, utilizing our method, we also develop an image-processing algorithm for simultaneous pose estimation based on visual features related to the head part of each agent. We conducted estimation experiments based on video recordings of cricket-robot interactions and confirmed that our proposed method could estimate the pose of both agents. This method will help to efficiently analyze behavioral data during interaction experiments.

Keywords: pose estimation, interaction behavior experiments, image processing, micro mobile robot, mobiligence

1. Introduction

The results of the biological research often provide valuable and indispensable knowledge of how animals behave adaptively in real-time. They not only show the animals’ flexibility, but also suggest how to design artifacts with adaptive functions. From this viewpoint, new biological findings are having a major impact on society. The ordinary biological approach usually involves the analysis of target animals and the behaviors at each organizational level of biological systems. Recently, a new trend in biological approaches has emerged that involves agent-based behavior experiments such as the introduction of an artificial agent into behavioral experimental setup. In contrast to ordinary biological experiments, this approach is novel in that experimental conditions are set and adjusted by operating the artificial agent such as a mobile robot. This kind of methodology has attracted attention in the field of experimental biology, as noted in a previous report on [1]. Behavioral experiments utilizing artificial agent are expected to provide new biological knowledge. To achieve this, experiments have to be conducted while sustaining objectivity and reproducibility. Such experiments can also contribute to a deeper understanding of the adaptive behavior selection mechanism of the cricket by expanding our approach to behavioral biology through the fusion of biology and engineering. Halloy et al. [2] conducted an experiment on the coexistence of cockroaches and autonomous micro mobile robots with a chemical substance on their cockroach-body surface. They showed that the number of individuals in a shelter affected the behavior of the cockroach group in an environment with no intervention of the operators. However, the interactions between individuals were controlled indirectly by adjusting the population density of individuals in the shelter. Since individual behavior and group behavior emerged from interactions between individuals, experimental systems
contributing to important discussions on interaction are imperative.

On the other hand, we have already started to develop an active interaction/on-line visual data gathering system for the cricket (*Gryllus bimaculatus*) [3]. The cricket is known as a solitary species of insect, but it shows behavior switching through the interaction with others that is representative of the pheromone behavior in insects [4]. The cricket also switches its tendency of behavior selection based on its own experiences or the context in the other situations [5]. Webb [6], [7] has conducted significant works on modeling the neural system of insects which the sensory-motor system was clarified on the basis of the knowledge of the insect’s behavior and response. Our motivation is to investigate the possibility of inducing behavioral switching through a physical interaction [8]. Therefore, we developed a system allowing a micro mobile robot to interact with crickets and we have conducted several interaction experiments utilizing this system. Our developed system was implemented to record the trajectories of both of a robot and a cricket by image processing. However, in order to analyze the behavior observed in such interaction experiments, it is necessary not only to track the positions of the agents but also to detect their poses. This is because, for example, in the fighting behavior between male crickets, the cricket generally shows behaviors such as antenna fencing, fighting, avoidance by the loser, chasing by the winner and so on. Therefore, the analysis of such behavior recorded on video is still hard work for experimenters, and attempts should be made to support the analysis of such recorded videos obtained by interaction experiments.

In this paper, we attempt to realize simultaneous pose estimation of a cricket and a robot from a video of their interaction taken during experiments. An image-processing algorithm for the pose estimation of both agents is presented and we conduct video analysis experiments. The results of our trials also confirm that our method works accurately.

2. Video Recording and Tracking of Agents

The configuration of the developed prototype system is shown in Fig. 1. The system consists of a control unit for the micro mobile robot and a measurement unit for on-line behavioral data collection. Command transmission to the robot, image capturing/processing and data storing can also be done simultaneously. In this system, we utilize a Macbook Pro (CPU: 2.6GHz Core 2 Duo, Memory: 4.0GB) and IRTrans (a product of IRTrans GmbH, Germany) for commanding the robot via infrared communication. A control pad with a USB plug is also utilized for maneuvering micro mobile robot.

A USB-connected camera (Logicool C905m) is utilized to capture images from the top view. This system uses Ecobe, a two-differential-wheel-driven micro mobile robot as an interacting physical agent [9], which has dimensions of 18×18×22 [mm].

To trigger the pheromone behavior of the cricket, the cuticular substances of the cricket must be coated on the robot’s body. Here, the head part of a cricket is placed on the fore of the body (Fig. 2, male cricket interacting with a micro mobile robot with a male cricket head). When the robot with a male cricket head was operated to interact with a male cricket, it was observed that the male cricket showed behavior toward the robot similar to that toward ordinary male crickets [10]. Interaction experiments were also conducted using a robot with a female head [11]. It was also found that the male cricket approached the robot and made a sound resembling a courtship song in response to the female-headed robot. Utilizing this system, we attempted fundamental interaction experiments (Fig. 3) and confirmed that video data could be gathered on-line [3].

Kai and Okada [12] developed a system for recording the locomotion and neural activities of the cricket. It included a visual tracking and pose estimation method for the cricket that utilized artificial markers added on the cricket’s body. However, such a system would be unnatural for the cricket. For behavior analysis, a situation as natural as possible is required in the experiment. Therefore, we developed an image-based tracking method for the cricket in a natural situation. To track both agents simultaneously, we performed the
color tracking of subjects similarly to [3], where the body surface color (blackish) of the cricket and a color marker (reddish) on the robot are tracked. In our developed tracking method, no markers or artificial objects are added on the cricket’s body. Here, to track the color feature in the image, a tracking algorithm with a particle filter is employed. In particular, we utilize the Condensation algorithm [13] and also the Mahalanobis distance for outlier analysis [14]. As a transition model of each particle, we utilize a randomized motion model for particle-filter-based tracking. The position of each agent (with blackish and reddish colors) is estimated simultaneously from the positions of the particles of each group. Okuda et al. [15] recently reported a multiple-crickets tracking method using image-processing, however, it only tracked the positions of the crickets.

In this paper, we discuss image-based pose estimation for both a cricket without marker and a robot with a marker from videos recorded with the above-mentioned system.

3. **Pose Estimation of Both Agents**

In this section, we describe the pose estimation algorithm utilizing image processing on a video captured by above-mentioned system. Basically, the following descriptions presume to use of our presented method [3], which provides the positions of both agents. Figure 3 shows an example of a birds-eye view image which was captured by the developed interaction system. The aim is to derive the pose of each agent (cricket and robot) from this type of image continuously and simultaneously.

We focused on the head part of each agent, which can be utilized as a visual feature and the estimated position of the agent. The key idea is to search for the head-part region in the area surrounding the estimated position using a particle-filter-based algorithm. This means that the vector from the agent’s position to the location of the head part can be considered as an estimated pose vector. The concrete derivation algorithm used in image processing is as follows:

1) Get an estimated robot’s position \( r_p \) and estimate the cricket’s position \( c_p \) by a particle filter tracking process (using the method and the parameters in [3]) on the captured image.

2) Select an inscribed square in the tracked particles determined from the Mahalanobis distance (using the method and the parameters in [3]). Here, \( r_q_i \) and \( c_q_j \) express the positions of a pixel in inscribed squares for the robot and the cricket, respectively.

3) Binarize the color values on all pixels with the threshold in each inscribed square and label them by following rule:

\[
\begin{align*}
  r_q_i & = \begin{cases} 
  1, & \text{pixel value of } r_q_i \geq r_q \
  0, & \text{otherwise}
  \end{cases} \\
  c_q_j & = \begin{cases} 
  1, & \text{pixel value of } c_q_j \geq c_q \
  0, & \text{otherwise}
  \end{cases}
\end{align*}
\]

where \( r_q \) and \( c_q \) are the threshold to (R,G,B) values of the pixels used in the binarization operation.

4) Perform a smoothing operation with the median filter and update the color value on each pixel as follows.

\[
\begin{align*}
  r_q_i & \leftarrow \text{median}( r_q_i) \\
  c_q_j & \leftarrow \text{median}( c_q_j)
\end{align*}
\]

where \( \text{median}(\cdot) \) is the median operator.

5) Calculate the average value of the blackish pixels in each inscribed square to estimate robot’s head position \( r_p_h \) and cricket’s head position \( c_p_h \).

\[
\begin{align*}
  r_p_h & = \frac{1}{\sum_{i=1}^{\omega} \sum_{l=1}^{M} r_q_i} \sum_{i=1}^{\omega} \sum_{l=1}^{M} r_q_i \ r_q_i \\
  c_p_h & = \frac{1}{\sum_{j=1}^{\omega} \sum_{l=1}^{M} c_q_j} \sum_{j=1}^{\omega} \sum_{l=1}^{M} c_q_j \ c_q_j
\end{align*}
\]

where \( \omega \) and \( M \) are the total numbers of the pixels in the inscribed squares for the robot and the cricket, respectively.

6) Estimate the mobile robot’s pose vector \( r_\eta \) and the pose vector of the cricket \( c_\eta \) as follows.

\[
\begin{align*}
  r_\eta & = r_p_h - r_p \\
  c_\eta & = c_p_h - c_p
\end{align*}
\]

Here, we can estimate the pose of each agent as the orientation of each derived vector.

7) Back to 1).

By utilizing this algorithm, blackish area in the surround area of the agent’s position is derived and pose estimation is done.

Figure 4 shows an example of a snapshot of this process showing the interaction between a male cricket and a robot with a male head part. The male cricket responded to the pheromones on the head part on the robot and showed aggressive behavior toward the robot. The upper right of this figure shows two small images used for head-part discrimination in this process. The left one is the binarized result for the cricket and the right one shows the result for the robot. The inscribed squares are adjustable and depend on the outlier analysis based on
the Mahalanobis distance used to evaluate the group of tracking particles [3]. In this case, the size of the inscribed square for the cricket was 52×52[pixels] and that for the cricket was 45×45[pixels]. Figure 5 shows magnifications of the images in the upper right of Fig. 4 and Fig. 6 shows the result of processing by the proposed algorithm. The particles used for tracking, the circles for the inscribed squares and the estimated pose vector are shown in Fig. 6. In practical processing, this algorithm is applied to each frame of a prerecorded video and the estimated poses are recorded as analyzed data.

4. Experiments

By utilizing the proposed algorithm, we conducted processing trials utilizing videos that were recorded during the interaction experiments between a male cricket and a mobile robot. Here, the size of the arena for the interaction experiment was 200×150 [mm]. The resolution of the captured images was 480×360 [pixels] and the frame rate of the video was 20 [fps]. In the step 4) of the algorithm, the median filter is used for the smoothing operation and the size of unit region of the median filter was 3×3[pixels]. The threshold \( \zeta \) was set as \((R,G,B) = (40,40,40)\) and \( \zeta \) was set as \((R,G,B) = (7,7,7)\). These thresholds were determined by trial and error from several images extracted from the videos. The micro mobile robot was commanded by the operator through the developed interaction system. Figure 7 shows another example of an image, showing the interaction between a male cricket and the robot with male head part. The right of Fig. 7 shows a magnified image and the particles used for tracking, the estimated vectors and the vector between the tracked positions of the agents.

To observe and analyze the cricket’s behavior, it is helpful to use the measured distance and estimated poses. For example, at the start of fighting behavior between male crickets, male cricket goes head-to-head with the opponent, then the crickets start fighting head-to-head when one cricket selects aggressive behavior. When the cricket selects avoidance behavior, its pose is changed to a different direction from that of the opponent. Therefore, the difference in the pose makes such evaluation easy. Figure 8 shows the results of each estimated pose in the arena (the horizontal axis of the image is set to 0 [rad]), the relative pose angle between the agents and the distance between the agents. From the relationship between the relative pose angle and relative distance, we can estimate the state of the agents. The pose and relative angle data in Fig. 8 are mapped from \(-\pi\) [rad] to \(+\pi\) [rad]. At the boundary between \(-\pi\) [rad] and \(+\pi\) [rad], data is switched from the top to the bottom or from the bottom to the top. However the data changes continuously. When the relative pose angle is being around 3.141 [rad] and the relative distance is with around 50 [pixels], the cricket should interact the robot or start to fight. When the relative pose angle becomes around 0.000 [rad] or 6.283[rad], there is no interaction or the cricket selects avoidance behavior. For example, we can observe fighting behavior from 8.00 [sec] to 12.00 [sec] and avoidance behavior from 13.00 [sec] to 17.00 [sec]. After 36.00 [sec], the cricket showed aggressive behavior until 47.00 [sec]. To validate these results, Fig. 9 shows top views of the original video for comparison with Fig. 8. Comparison between the figures
confirmed that the algorithm can contribute to behavior analysis in cricket-robot interaction experiments.

5. Conclusion

In this paper, we described a pose estimation method for a cricket and a mobile robot based on image processing. The pose is estimated by tracking positions using a particle filter and detecting the head part in the region surrounding each agent. The experimental results, especially changes the estimated poses with time, are shown. Moreover, the estimated results were compared with top views of the robot and cricket in the arena. The results show that the proposed pose estimation process has sufficient performance to contribute to supporting the behavior observation tasks.

On the basis of these results and those in our previous works, we are currently developing a new systematic tool for ethological experiments. In particular, the aim is to maintain the reproducibility and tendencies observed in interaction experiments and also to evaluate behavior models using a physical agent. This tool will be an on-line automatic controlled interaction framework based on trajectory and pose estimation by using a micro mobile robot that exhibits programmed behaviors.
Acknowledgement

This work has been partially supported by KAKENHI (No. 25540118 and No. 23300113) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, the Cooperative Research Program of "Network Joint Research Center for Materials and Devices" and also NSFC project 90920301 in China.

References


Kuniaki Kawabata received his Ph.D. degree in electrical engineering from Hosei University in 1997. He worked at Biochemical Systems Lab. of RIKEN (The Institute of Physical and Chemical Research) as a special postdoctoral researcher from 1997 to 2000. In 2000, he joined the Advanced Engineering Center of RIKEN as a research scientist. In 2002, he joined the Distributed Adaptive Robotics Research Unit, RIKEN. In 2005, he became Unit Leader of the Distributed Adaptive Robotics Research Unit, RIKEN. In 2007, he became Unit Leader of the Intelligent System Research Unit of RIKEN. In 2011, he became Team Leader of the RIKEN-XJTU Joint Research Team and is currently Unit Leader of the RIKEN-XJTU Joint Research Unit. His research interests cover distributed autonomous systems, networked systems, the dynamic modeling of social adaptability and mobile robots. He is a member of RSJ, SICE, JSME and IEEE.

Hitoshi Aonuma received his D.Sc. degree from the Faculty of Science, Hokkaido University, Japan, in 1998. From 1995 to 1996, he was with the Graduate School of Science, Hokkaido University, as a JSPS Research Fellow. In 1998, he was a Research Associate for BBSRC, School of Biological Sciences, University of Southampton, U.K., where he was a JSPS Research Fellow from 1999 to 2000. From 2001 to 2003, he was an Assistant Professor, at the Research Institute for Electronic Science, Hokkaido University, where he has been an Associate Professor since 2003. His research interest includes neuroethology, neurobiology and animal physiology. He is a member of the International Society for Neuroethology, International Society for Invertebrate Neurobiology, Zoological Society of Japan, Japanese Society for Comparative Physiology and Biochemistry and Society of Evolutionary Studies.

Satoru Takahashi received his Ph.D. degree from Tokyo Denki University in 2001. From 1991 to 1993 he was a Researcher of the Nano Technology Research Center, Canon Inc., Japan. From 1997 to 1999 he served as an Instructor in the Department of Information Sciences, Tokyo Denki University. From 1999 to 2000, he was a Visiting Researcher in the Systems, Science and Mathematics
Department at Washington University in St. Louis. In 2001, he served as a Research Assistant in the Mathematics and Statistics Department at Texas Tech University. Since 2003, he has been an Associate Professor in the Intelligent Mechanical Systems Engineering Department at Kagawa University. In 2012, he joined the RIKEN-XJTU Joint Research Team at RIKEN (currently Research Unit) as a Visiting Scientist. His research interests include bioimage informatics, robotics and mechatronics. He is a member of RSJ, SICE, ISCIE, IEEJ, JSPE and IEEE.

Koh Hosoda received his Ph.D. degree in mechanical engineering from Kyoto University, Japan in 1993. He was an Assistant Professor at the Mechanical Engineering Department from 1993 to 1997, and an Associate Professor at the Graduate School of Engineering from 1997 to 2010, at Osaka University. He was a Guest Professor at the Artificial Intelligence Laboratory, University of Zurich from Apr. 1998 to Mar. 1999. He was a Group Leader of JST Asada ERATO Project from 2005 to 2010. From 2010 to 2014, he was a Professor of Graduate School of Information Science and Technology, Osaka University. Since 2014, he has been a Professor of Graduate School of Engineering Science, Osaka University.

Jianru Xue received his Ph.D. degree from Xi’an Jiaotong University in 2003. He worked at Fuji Xerox Co., Ltd. from 2002 to 2003 and was a visiting researcher at the University of California, Los Angeles, from 2008 to 2009. He is now a Professor of institute of artificial intelligence and robotics, Xi’an Jiaotong University. His research interests include computer vision, visual navigation, and video coding based on analysis. He has published more than 40 papers in top journals and the conferences including IEEE Transactions on Image Processing, IEEE Transactions on System, Man and Cybernetics Part B, ICCV and ACM MM.

(Received January 8, 2014; revised March 19, 2014)