2021 The 21st International Conference on Control, Automation and Systems (ICCAS 2021)
Ramada Plaza Hotel, Jeju, Korea, Oct. 12∼15, 2021

Actualization of Deep Ego-motion Classification on Miniaturized Octagonal Compound Eye Camera

Hwiyeon Yoo[1*], Jungho Yi[1*], Jong Mo Seo[1], and Songhwai Oh[1**]

[1]Department of Electrical and Computer Engineering, Seoul National University, Seoul, 13391, Korea
(hwiyeon.yoo@rllab.snu.ac.kr,myrmidon91@snu.ac.kr, callme@snu.ac.kr, songhwai@snu.ac.kr)
* Equal contribution ** Corresponding author

Abstract: Vision system in robotics is growing faster than ever. While a majority of studies rely on single-lens cameras in obtaining visual information from the surrounding environment, a compound eye camera that imitates the eyes of arthropods and insects has shown its potential in previous researches. They have used several simulations of the compound eye images in an ideal, refined, and computerized circumstance in spite of the absence of a practical device that offers real-world data due to its limited visible range of a few centimeters. On the one hand, micro-lens arrays that mimic compound eyes have been conducted through various fabrication processes and forms without providing definitive post-processed applications with an appropriate algorithm that can handle such information. In this research, the deep ego-motion classification algorithm and octagonal compound eye camera that has n_l=3 are harmonized for viable example of the compound eye vision system.

Keywords: Deep Ego-motion Classification, Neural Network, Compound Eye, Micro-lens Array, Computer Vision, Biomimetics, Software-Hardware Integration

1. INTRODUCTION

A compound eye camera is a sophisticated vision structure that mimics an insect eye. It consists of a number of the single eye that obtains a small image from different aspects from one another. By simultaneously capturing images through single eyes, a compound eye camera enables us to achieve a larger field-of-view (FOV) compared to ordinary cameras [1]. By virtue of the recent success of the deep-neural-network-based computer vision algorithms, there have been researches to develop computer vision algorithms further that suit the unique structure of the compound eye camera using deep neural network [2–4]. Despite their achievements in developing algorithms, these previous works have a limitation in that the experiments were mainly done for the simulated compound eye images assuming ideal condition due to the absence of real, designated compound eye camera. In this paper, we focus on one of these algorithms, deep ego-motion classification [4], and actualize it on a real-world device called a miniaturized octagonal compound eye camera. The results show that the integration of the deep ego-motion algorithm and the proposed hardware work seamlessly, which implies we successfully overcome the gap between the ideal simulation and the real-world situation. In addition, it has verified that the core part of the deep ego-motion classification algorithm is the unique structure of the compound eye camera, amongst other parameters .

2. METHOD

2.1 Deep Ego-motion Classification

The algorithm mainly consists of two parts: the local motion classification network and the aggregation of local classifications. The local motion classification, M, is a convolutional neural network (CNN) whose input is two vectorized compound images and output is single-eye-wise motion classifications. Here, a vectorized compound image $I^c \in \mathbb{R}^{n_l \times n_l \times 3S^2}$ is a transformed compound image that each single eye RGB image of $S \times S$ pixels is vectorized to $\mathbb{R}^{3S^2}$, where there exist $n_l \times n_l$ number of single eyes. The output single-eye-wise motion classifications represents the estimated discrete ego-motion with respect to each single eye observation ($M(I^c_n, I^c_m) \in \mathbb{R}^{n_l \times n_l \times n_d}$, where n_d is the number of discrete ego-motion candidates).

The aggregation of local classifications is voting step to determine the final classification which is formulated as following, $Class_f = argmax \sum_{i,j} M(I^c_n, I^c_m)_{i,j}$, where $Class_f$ is the result of the final classification and $M(I^c_n, I^c_m)_{i,j}$ is a local classification result of a single eye on i-th row and j-th column of the compound eye [4].

2.2 Micro-Lens Array

Polydimethylsiloxane (PDMS) and Polytetrafluoroethylene (PTFE) are used as materials for the lenses and the light-screen, respectively. Unlike ordinary compound eye mimicking micro-lens arrays with either rectangular or hexagonal shape in terms of its arrangements, the octagonal compound eye modified for this integration has eight lenses surrounding the one in the center for better suitability. In the case of rectangular alignment, the azimuth angle of the lenses varies, demanding complicated mathematical design for post-processing. The hexagonal alignment offers a concise azimuth angle compared to the rectangular, but it is not appropriate to be used for the existing ego-motion classification algorithm which requires $n_l \times n_l$ formation of eyes. The micro-lens array is not free from spherical aberration. To minimize the distortions on

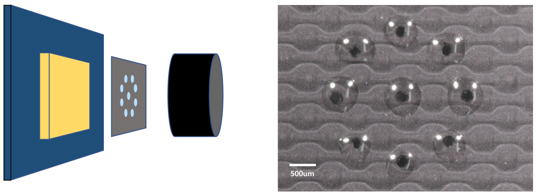

Fig. 1. Schematics of the camera setup (left) image sensor, micro-lens array, and the aperture, from left to right. (right) The fabricated PDMS Micro-Lens Array on PTFE light blocking layer, each of which has 600 μm diameter, equally separated by 1.5 mm.

the edge of the image, a cylindrical aperture is used as shown in Fig. 1. This also enables us to precisely determine the exact boundary of the image obtained. Then, 210×210 rectangular sample around the center of each image is extracted to be processed by the algorithm.

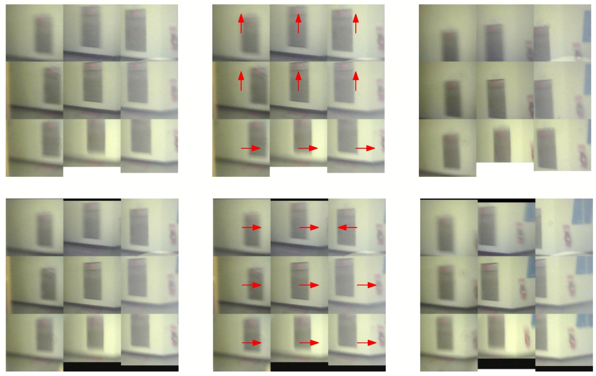

Fig. 2. Successful examples of the deep ego-motion classification on the miniaturized octagonal compound eye camera. (left) Input image 1 which is captured before movement, (center) local motion classification on image 1, (right) input image 2 which is captured after movement.

3. EXPERIMENTS AND RESULTS

In this paper, we apply the deep ego-motion classification to the miniaturized octagonal compound eye camera proposed in Section 2.2. Since each single eye image from the camera is in circular form, 210×210 rectangular image is cropped from the center of each single eye and vectorized to fit the form of I_c ($S = 210$). Also, the proposed compound eye camera has 2 layers of single eyes as shown in Fig 1(right), so n_l becomes 3. The deep ego-motion classification network is trained from scratch with these parameters to obtain a model that suits the proposed hardware. For training and evaluation, the data set from [4] is used and the task is simplified for the classification of 4 discrete ego-motion directions. The trained network achieved 86.91% accuracy on the evaluation data set. The results of using observations from the miniaturized octagonal compound eye camera is in Fig. 2. Considering the voting step, it is clear that both examples successfully classified the ego-motion of the camera.

4. CONCLUSION AND DISCUSSION

In this paper, the deep ego-motion classification is actualized on a miniaturized octagonal compound eye camera with $n_l = 3$. Although the observations from the proposed compound eye camera have differences with the ideal compound images from a simulation with larger n_l, the deep ego-motion classification algorithm shows feasible results with small n_l. However, a dilemma remains when it comes to the size of each micro-lens. In order to see further, the diameter of the lens must be increased. Otherwise, only a few centimeters from the lenses would be the visible range. As the diameter of the lens increases, the larger the resulting image gets, which limits the number of images cast on the image sensor, making it impossible to cast more than 3×3 images (n_l=3) at once. The problem can be relieved by adopting a larger image sensor, or by imitating apposition compound eye type which has designated image sensor for each micro-lens. In the future, combinations of a micro-lens array with further visible range and a more powerful neural network based algorithm for different applications would be discussed.

ACKNOWLEDGEMENT

This research was supported by a grant to Bio-Mimetic Robot Research Center Funded by Defense Acquisition Program Administration, and by Agency for Defense Development (UD190018ID).

REFERENCES

[1] M. Seo, J.-M. Seo, D.-I. Cho, and K.-I. Koo, “Insect-mimetic imaging system based on a microlens array fabricated by a patterned-layer integrating soft lithography process,” *Sensors*, vol. 18, no. 7, p. 2011, 2018.

[2] H. Yoo, D. Lee, G. Cha, and S. Oh, “Estimating objectness using a compound eye camera,” in *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2017.

[3] G. Cha, H. Yoo, D. Lee, and S. Oh, “Light-weight semantic segmentation for compound images,” in *Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2017.

[4] H. Yoo, G. Cha, and S. Oh, “Deep ego-motion classifiers for compound eye cameras,” *Sensors*, vol. 19, no. 23, p. 5275, 2019.