Proceedings of ACML-TAB2019

Flow Histogram-based Recurrent Neural Network for Visual Odometry Estimation

Yukihiro Achiha YUKIHIRO@MPRG.CS.CHUBU.AC.JP
Chubu University

Tsubasa Hirakawa HIRAKAWA@MPRG.CS.CHUBU.AC.JP
Chubu University

Takayoshi Yamashita TAKAYOSHI@ISC.CHUBU.AC.JP
Chubu University

Hironobu Fujiyoshi FUJIYOSHI@ISC.CHUBU.AC.JP
Chubu University

Abstract

Developing driving assistance or autonomous driving systems requires to understand the vehicle position and motion. To classify the vehicle movement categories from on-board camera images, convolutional recurrent neural network (C-RNN) has been developed. However, it is difficult to estimate odometry precisely. In this paper, we propose a visual odometry estimation method by extending C-RNN framework. The proposed method utilizes histograms of optical flow values to achieve robust odometry estimation for noise of estimated optical flow or difference of scenes. Our experimental results show that the proposed method can estimate odometry.

Keywords: Visual odometry estimation, Convolutional recurrent neural network, Flow histogram

1. Introduction

Odorimetry estimation is a fundamental problem on developing autonomous driving and driving assistance systems. While a number of approaches have been developed such as the use of GPS, gyro sensors or in-vehicle camera images. Among them, visual odometry estimation that utilizes camera images has a much potentinal because visual odometry estimation does not require any additional sensors and is robust for weather or building environtments (e.g., tunnels) unlike GPS. As a method related to visual odometry estimation, a convolutional recurrent neural network (C-RNN) (Zuo et al., 2015) can be used to consider tempral changes from images and to classify vehicle movement class from optical flow. However, this method classifies discrete categories of movement and does not estimate odometry as continous values.

In this paper, we propose a method for visual odometry esitmation from in-vehicle camera images. The propsoed method is based on a recurrent neural network (RNN) architecture. We use optical flow data as an input for the RNN. Specifically, we divide optical flow data into seral grids and make histograms with respect to each grid region. This histogram representation is robust for estimation noise of optical flow or diffenrence of the esitmated scenes. Experimental results shows that our approach achives higher estimation accuracies.

2. Method

The prposed method is based on the optical flow and RNN architecture. First, we extract optical flow from camera images. To estimate optical flow, we applied TV-L1 algorithm. The extracted optical flow is divided into several grids and flow histograms are created for each grids. To divide optical flow data, we use a couple of division approachs based on SIFT (Lowe, 1999) and GLOH (Mikolajczyk and Schmid, 2003) that are local image features. The created histograms are input to a long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1997). In our experiment, we compare the estimation accuracies between the two division approaches.

3. Results and Conclusions

We evaluated the proposed method with in-vehicle camera image datasets. By using this dataset, we estimate the following three values: yaw, pitch, and migration length. As results, SIFT-like grid outperforms GLOH-like grid. Moreover, we compare the propsoed method with a conventiolan C-RNN architecure. The result shows that the accuracy of the prposed method is better than that of C-RNN. However, the esimation error of migration length is rather large than yaw and pitch. Therefore, the improvement of estimation accuracy on migration length is one of our future work.

Acknowledgments

This work was supported in part by JSPS KAKENHI grant number JP16H06540.

References

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

D. G. Lowe. Object recognition from local scale-invariant features. In *IEEE International Conference on Computer Vision (ICCV)*, 1999.

K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.

Z. Zuo, B. Shuai, G. Wang, X. Liu, X. Wang, B. Wang, and Y. Chen. Convolutional recurrent neural networks: Learning spatial dependencies for image representation. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2015.