스트레오 정합을 위한 가이드 영상이 필요 없는 심층 집합비용 기법

*윌리엄, 박인규
인하대학교 정보통신공학과
e-mail : 22112195@inha.edu, pik@inha.ac.kr

Deep Cost Aggregation without A Guidance Image for Stereo Matching

*Williem, In Kyu Park
Department of Information and Communication Engineering
Inha University

Abstract

This paper presents a deep self-guided cost aggregation method for obtaining accurate disparity map from a pair of stereo image. Compared to conventional cost aggregation methods, the proposed method does not require any guidance image. Deep learning approach is utilized for performing self-guided cost aggregation. The proposed deep learning network consists of two sub-networks including dynamic weight network and descending filtering network. Experimental results show that the proposed method achieves better results even though it does not employ any guidance image.

I. Introduction

Cost aggregation plays an important role for local or global stereo matching. Most of the methods share similar concept with joint image filtering that utilizes a guidance image. Thus, the existence of a guidance image is inevitable. They assume that the weights of corresponding pixels on color image are similar with the weights on disparity image.

However, this assumption often fails when a patch with high texture variance has similar disparity value. In this paper, we propose a novel cost aggregation method that removes the dependency on the guidance image on the fly. A deep convolutional neural network is utilized in our approach so that a cost volume slice can perform a self-guided cost aggregation. To our best knowledge, the proposed work is the first self-guided cost aggregation method based on deep learning approach.

II. Proposed Method

2.1 Deep Neural Network Architecture

The proposed deep cost aggregation network consits of two sub-networks, which are dynamic weight network and descending filtering network. The proposed dynamic weight network generates a set of dynamic weights for each input pixel. These weights are useful as the guidance for the filtering network. The descending filtering network performs an edge-aware filtering operation for the input cost volume slice. Thus, we utilize an edge-aware filtered cost volume slice as the learning target. The overview of the network architecture is shown in

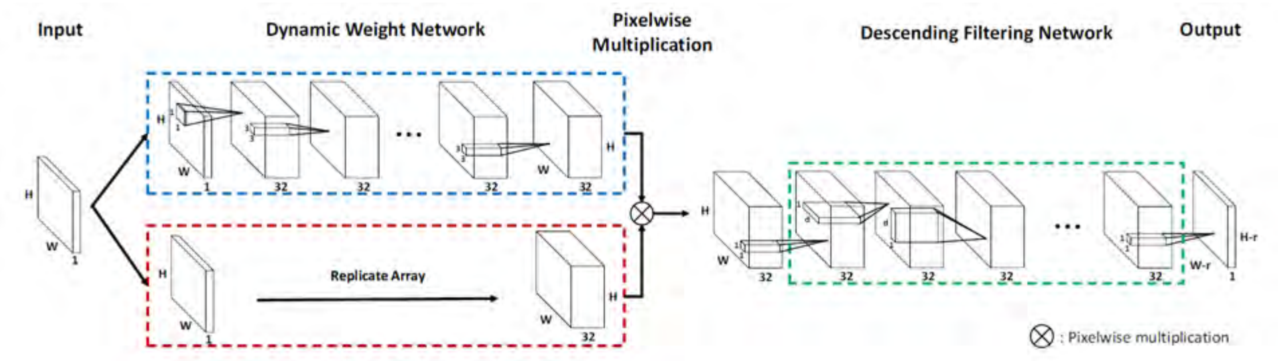

Figure 1 Pipeline of the proposed deep self-guided cost aggregation network.

Figure 1 where W and H are the width and height of the input. d and r are the diameter and radius of a support region, respectively. The loss function consists of feature reconstruction loss and mean square loss functions.

2.2 Dataset Generation

Due to the absence of the ground truth cost volume slice, it is difficult to apply deep learning approach on the cost aggregation step. Thus, we utilize a guided filtered cost volume as the target. The ground truth disparity map is utilized as the guidance. In this paper, we randomly extract 1000 patches from each training image where 80% of the patches are thos around the depth discontinuities. Data augmentation is performed to consider the rotation and contrast invariance.

Ⅲ. Experimental Results

In this paper, we perform the evaluation using Middlebury dataset and KITTI dataset to validate the performance. Table 1 shows the comparison of average bad pixel percentage of 200 KITTI training images. Same evaluation metric with KITTI benchmark is utilized. It shows that the proposed method achieves the smallest average error. Figure 2 shows the qualitative result of KITTI dataset.

	GF	NL	DEEP
Non-Occluded Region	11.40	13.14	9.69
Whole Region	12.96	14.65	11.23

Table 1 Average error of KITTI training data

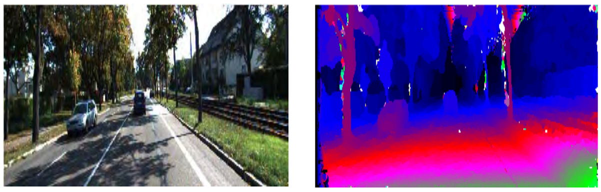

Figure 2 2015 KITTI training data. (Left) Input left image; (Right) Disparity map of the proposed method

Ⅳ. Conclusion

We proposed a novel deep convolutional network to perform a self-guided cost aggregation for stereo matching. We utilized a deep learning approach which consisted of two sub-networks including dynamic weight network and descending filtering network. Experiment results confirmed that the proposed method achieved the best results among the state-of-the-art techniques.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2016R1A2B4014731).

References

[1] A. Hosni et al.,“Fast cost-volume filtering for visual correspondence and beyond,” IEEE TPAMI, vol. 35, no. 2, pp. 504-511, 2013..

[2] Q. Yang, “A non-local cost aggregation method for stereo matching,” in Proc. of IEEECVPR, pp. 1402-1409, 2012.