Suppressing Features That Contain Disparity Edge for Stereo Matching

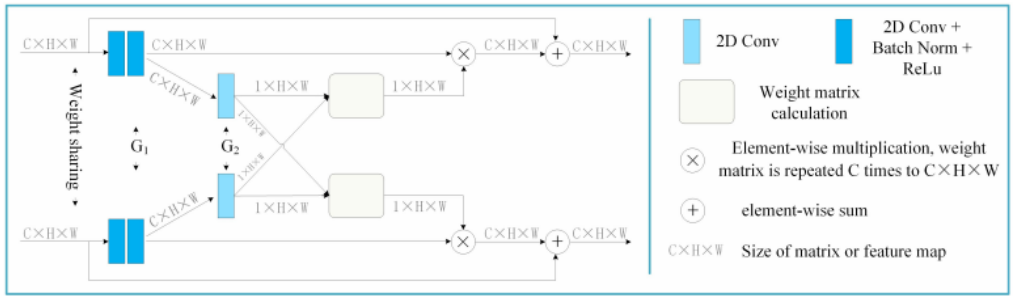
Xindong Ai[1], Zuliu Yang[1], Weida Yang[1], Yong Zhao[1*], Zhengzhong Yu[2] and Fuchi Li[2]
[1]Shenzhen Graduate School of Peking University (PKUSZ), [2]Shenzhen Apical Technology Co.

ABSTRACT

Existing networks for stereo matching usually use 2-D CNN as the feature extractor. However, objects are usually continuous in spatial, if an extracted feature contains disparity edge (the representation of this feature on original image contains disparity edge), then this feature usually not occur inside the region of an object. We propose a novel attention mechanism to suppress features containing disparity edge, named SDE-Attention (SDEA). We notice that features containing disparity edge are usually continuous in one image and discontinuous in another, which means that they usually have a greater difference in two feature maps of the same layer than features that don't contain disparity edge. SDEA calculate the weight matrix of the intermediate feature map according to this trait, then the weight matrix is multiplied to the intermediate feature map. experimental results show that our method has a significant improvement in accuracy.

BLOCK ARCHITECTURE

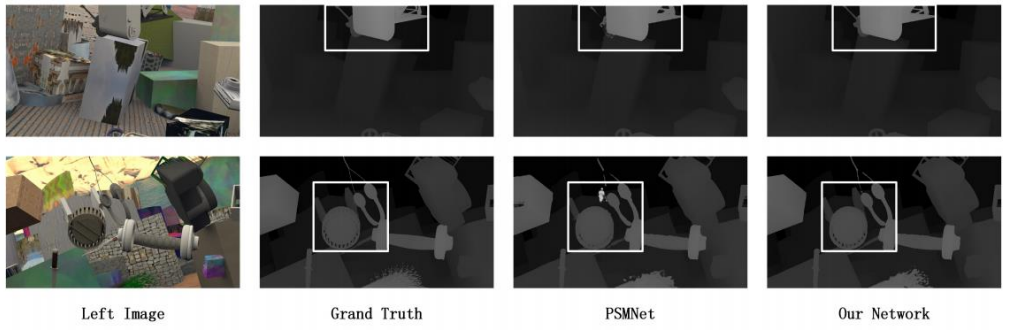
We proposed a novel attention mechanism (SDEA) and built it into a block (named SDEA-Block) to apply to CNN. Compared with ResBlock[1], SDEA-block adds almost negligible parameters and adds little forward time.

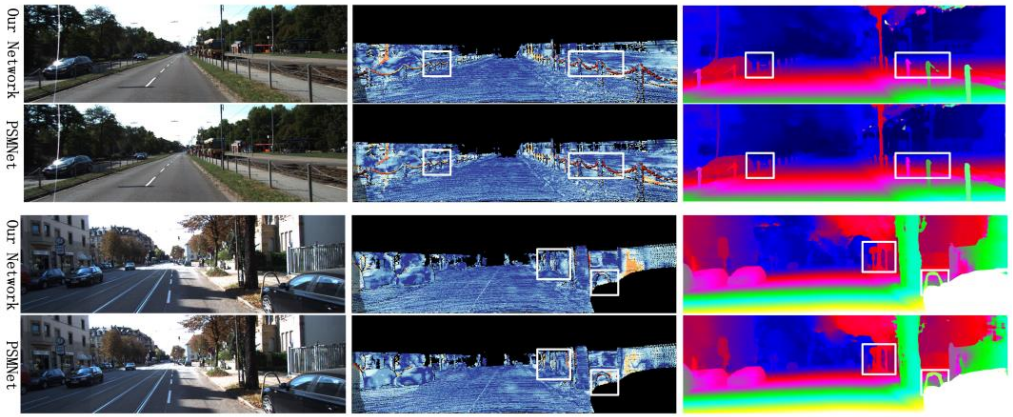


RESULTS

We apply SDEA-Block to PSMNet[2] and test it on SceneFlow[3] dataset and KITTI2015[4].

1. The visualization pictures:





2. The experimental results:

KITTI2015 RESULT

Model	All (%)			Noc (%)		
	D1-bg	D1-fg	D1-all	D1-bg	D1-fg	D1-all
GC-Net [20]	2.21	6.16	2.87	2.02	5.58	2.61
iResNet-i2e2 [19]	2.14	**3.45**	2.36	1.94	3.20	2.15
CRL [18]	2.48	3.59	2.67	2.32	**3.12**	2.45
SegStereo [22]	1.88	4.07	2.25	1.76	3.70	2.08
MCUA [5]	**1.69**	4.38	2.14	**1.55**	3.90	1.93
PSMNet [2]	1.86	4.62	2.32	1.71	4.31	2.14
SDEA-Net	1.71	4.17	**2.12**	1.56	3.76	1.93

SCENEFLOW DATASET RESULT

Model	EPE	Model	EPE
SDEA-Net	**0.77**	GwcNet-g [4]	0.79
PSMNet [2]	1.09	StereoNet [23]	1.10
CRL [18]	1.32	SegStereo [22]	1.45

CONCLUSION

In In this paper, we propose a general attention block for stereo matching, namely SDEA-Block, which aims at suppressing the features containing disparity edge. For the two given feature maps obtained by the input through two 3 × 3 convolution layers, SDEA-Block uses one 1 × 1 convolutional layer to aggregate information and reduce their channel to 1. For all points in each feature map with dimension (channel) 1, SDEA-Block searches for the points with the minimum difference, which is in a specific range of the corresponding feature map, then calculate the weight matrix of the two given feature maps based on this minimum difference, and the smaller this minimum difference means the greater the calculated weight. Experimental results demonstrate the effectiveness of SDEA-Block.

REFERENCE

[1] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2016) 770–778.

[2] Chang, J.R., Chen, Y.S.: Pyramid stereo matching network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018) 5410–5418.

[3] Mayer, N., Ilg, E., Hausser, P., Fischer, P., Cremers, D., Dosovitskiy, A., Brox, T.: A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2016) 4040– 4048.

[4] Menze, M., Geiger, A.: Object scene flow for autonomous vehicles. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2015) 3061–3070.