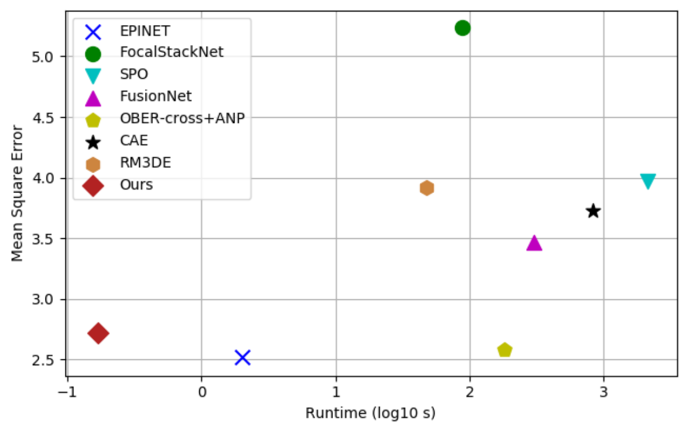
IEEE International Conference on Pattern Recognition(ICPR) 2020

Fast and Efficient Network for Field Disparity Estimation

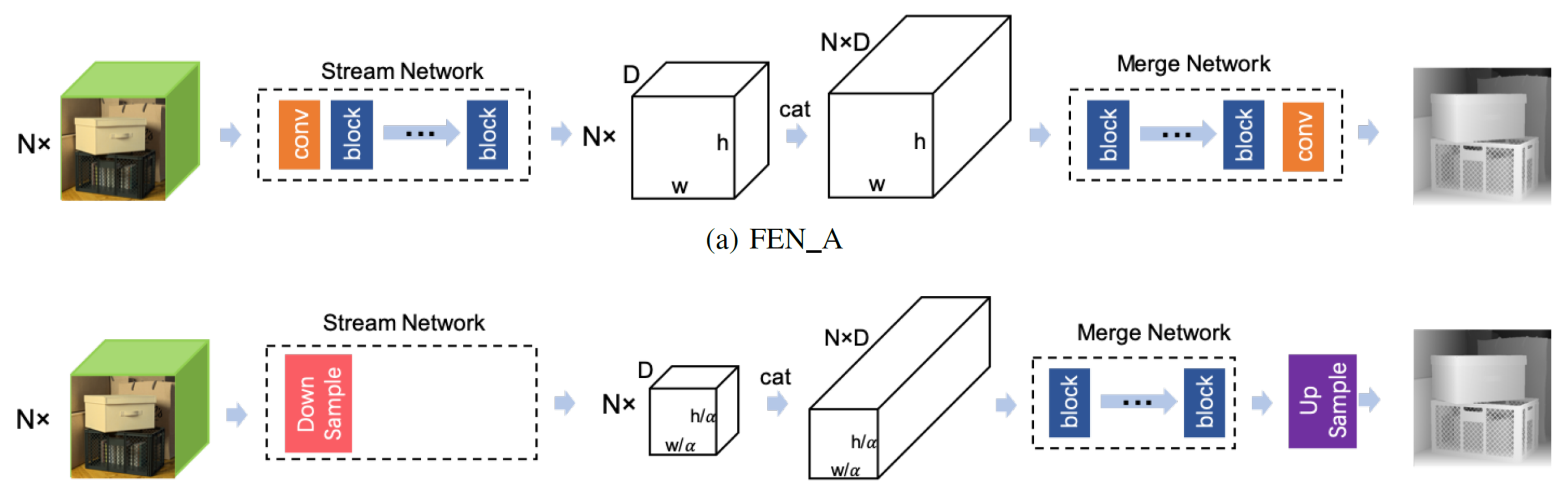
Dizhi Ma, Andrew Lumsdaine
University of Washington

Introduction



As with many imaging tasks, disparity estimation for light fields seems to be well-matched to machine learning approaches. Neural network methods can achieve the best result on the 4D light field benchmark dataset, continued effort to improve accuracy is resulting in diminishing returns. On the other hand, due to the growing importance of mobile and embedded devices, improving the efficiency is emerging as an important problem. In this paper, we improve the efficiency of existing neural net approaches for light field disparity estimation by introducing efficient network blocks, pruning redundant sections of the network and down-scale the resolution of feature vector. To improve performance, we also propose densely sampled epipolar image plane (EPI) volumes as input. Experiment results show that our approach can achieve similar results compared with SOTA methods while using only one-tenth runtime.

Methodology



(a) FEN_A

(b) FEN_B

We proposed a fast and efficient network with two-stage structure. The first stage extract the feature from EPI volumes, then we concatenate these features and pass it to the second part to predict the results. We used following ways to improve the efficiency and reduce the runtime of the model:

1. Efficient block. MobileNet shows that separable convolution kernel can also achieve good performance with many fewer parameters. We modified the inverted residuals module form MobileNetV2 for use as our basic block.

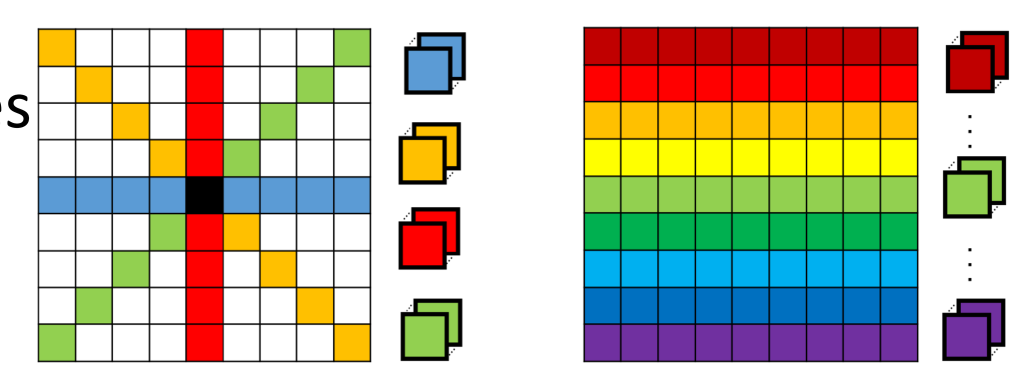
TABLE IV
COMPARISON OF SIZE AND COMPUTATION COST.

	# Params	MAdds
EPINET	5.1M	2.6T
FEN_A	1.5M	1.6T
FEN_B_×2	1.3M	175.0G
FEN_B_×4	1.4M	52.5G

2. Pruned stream network. Due to the input we used, the computation cost of this part will be multiply by the number of EPI volumes, which make it a considerable quantity. To reduce the computation of the stream network, we remove all the bottlenecks, only leave it with convolution kernel (s).

3. Down-scale feature. To further reduce the computation cost, we downsample the input with the stream network by simply setup the stride as 2.

We also introduce so called dense EPI volumes to make use of all the views and improve the performance.



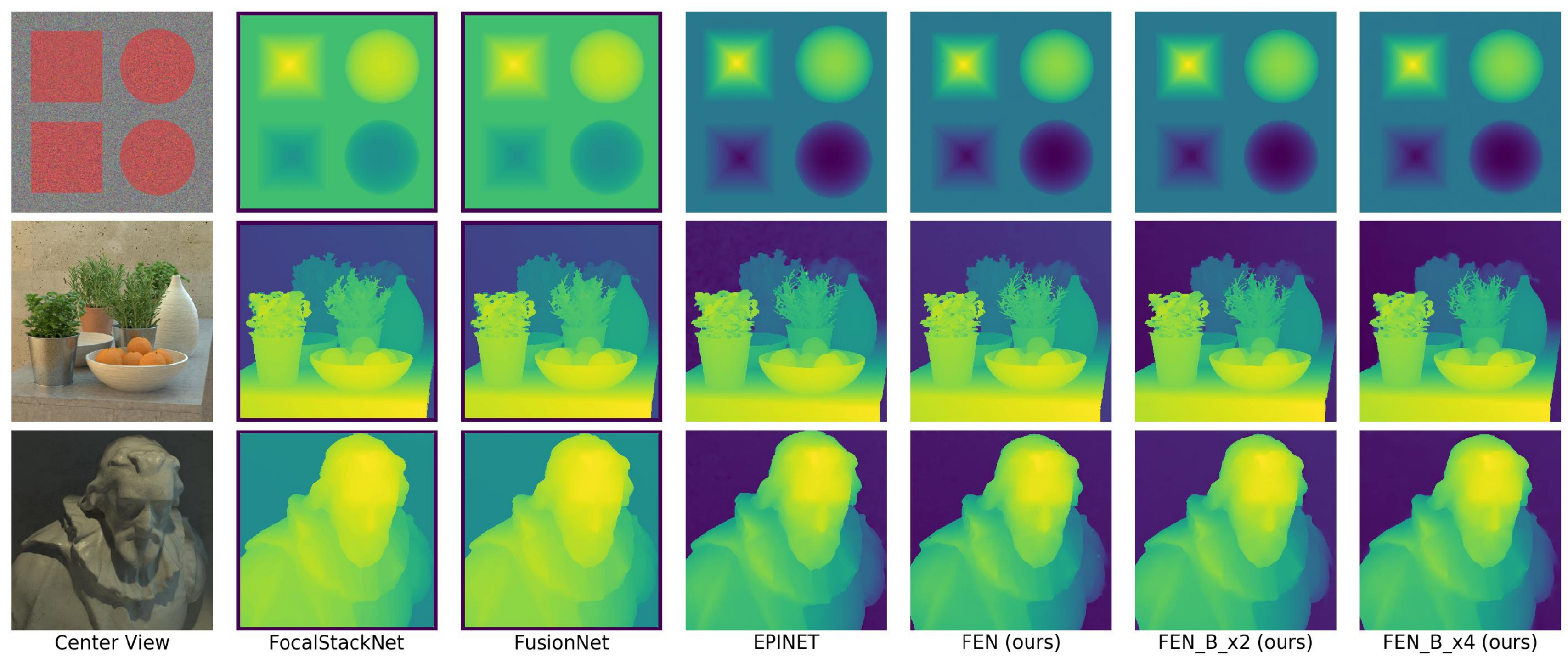
(a) 4 EPI input (b) 9 EPI input

Results

We compared the average bad pixel rate (with 0.07 threshold), mean square error (×100) and runtime metric with the SOTA methods.

method	BP0.07	MSE	Runtime	GPU
SPO-MO	5.71	3.52	4303.33	-
OBER-cross+ANP	**4.59**	2.58	182.99	-
FocalStackNet	5.26	5.24	88.19	Titan X
FusionNet	4.67	3.47	303.51	Titan X
EPINET	5.41	2.52	2.04	1080 Ti
FEN_A (ours)	6.03	**2.19**	2.00	
FEN_B_×2 (ours)	6.84	2.72	0.17	1070 Ti
FEN_B_×4 (ours)	9.18	3.17	**0.06**	

The FocalStackNet adopt refocus images generated with the light field images from the different focal plane. The FusionNet fuse hybrid cues include EPI and focal stack. Both FocalStackNet and FusionNet are the pixel-wise methods, so their runtimes are slower. With much fewer parameters and downsample strategy, our methods is much faster compare with the other method, but only slightly inferior in bad pixel rate and mean square error. Notice that We used Nvidia 1070 Ti GPU to benchmark the runtime of our methods. In the figure below, We also provide the visual result comparison. Although our quantitative values are slightly worse, visually, the proposed methods provide comparable results.



Conclusion and Acknowledge

In this paper, we propose a two-stage fast and efficient neural network for light field disparity estimation. Four techniques are applied to reduce the computation cost and improve the efficiency. First, we adopt the lite block approach to reduce the overall size of the model. Second, we prune the first stage of the network to reduce the computation cost of that stage. Third, we downsample the features to reduce the computation cost of the second stage. Finally, we also introduce a novel EPI volume input to improve the performance. Experimental results show our model can achieve similar performance compared with state-of-the-art methods with significantly improved run time.

Computational resources for this work were provided in partby the US Department of Energy, Office for Advanced Scientific Computing under No. 66150: "CENATE: The Center for Advanced Technology Evaluation" and the Laboratory Directed Research and Development program at PNNL under contract No. ND8577.

Contact: {dizhim, al75}@uw.edu