Deep Video Deblurring for Hand-held Cameras

UBC Adobe KAUST Universidad de la República Uruguay Duke University

Shuochen Su Mauricio Delbracio Jue Wang Guillermo Sapiro Wolfgang Heidrich Oliver Wang

CVPR July 21-26 Honolulu 2017

Introduction

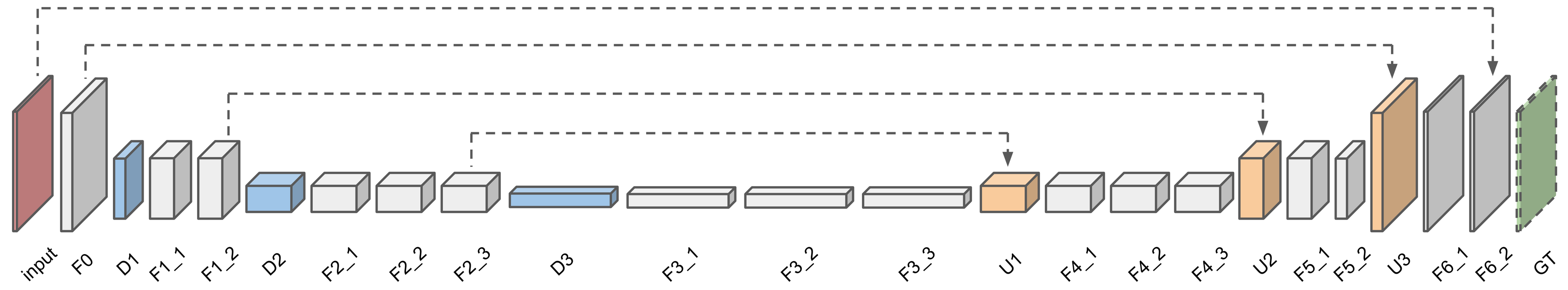
Problem: Motion blur from camera shake is a major problem in videos captured by hand-held devices. Traditionally, video-based approaches can take advantage of the abundant information across neighboring frames.

Challenge: Multi-frame alignment is a computationally expensive and fragile procedure. Methods that aggregate information must be able to identify warping artifacts from true contents, a task that requires high-level scene understanding.

Proposal: We introduce a deep learning solution to video deblurring, where a CNN is trained end-to-end to learn to accumulate information across frames. To train this network, we collect a dataset of real videos recorded with high frame rate cameras, and generate synthetic motion blur for supervision.

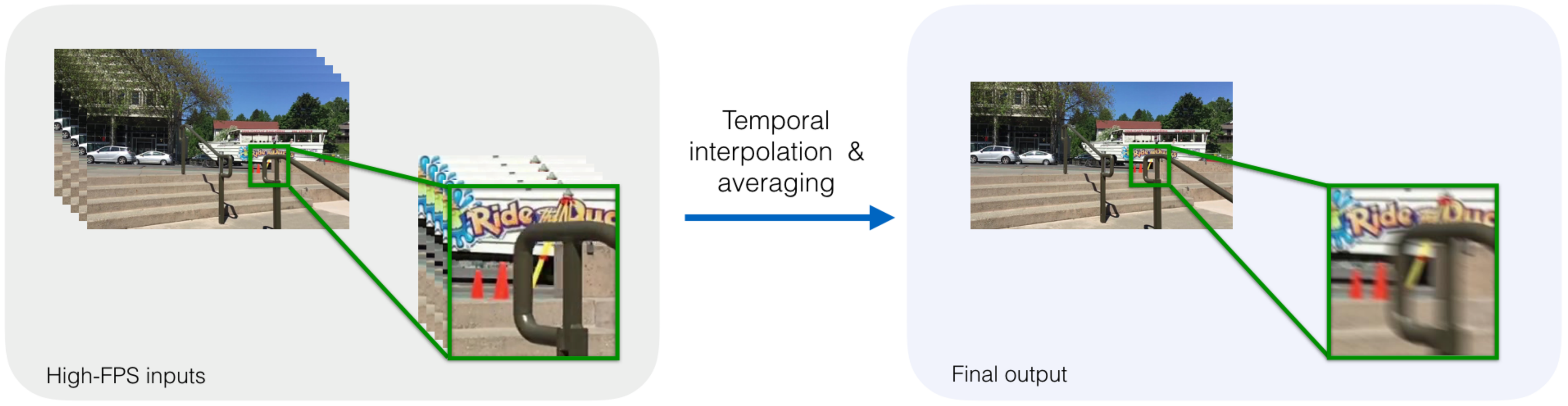
Approach

Network: An encoder-decoder network with symmetric skip connections is used, which increases the receptive fields and is yet easy to train.



The proposed DeblurNet architecture.

Realistic dataset: We collect real-world sharp videos at high frame rate (240fps), and synthetically create blurred ones by accumulating a number of short exposures to approximate a longer exposure.



Schematic of dataset generation.

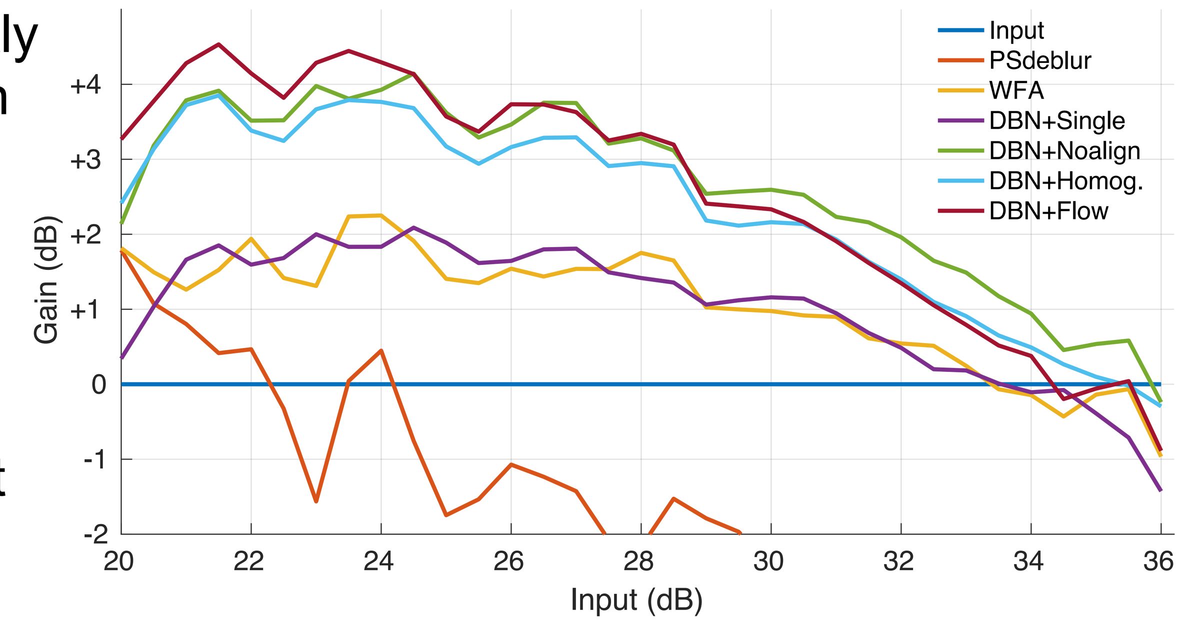
Experiments

Dataset and Training: In total, 71 videos are collected, each with 3-5s running time, to generate 6708 synthetic blurry frames with ground truth. These frames are subsequently augmented with flipping, rotating and scaling, from which we crop 2,146,560 128×128 random patches. We used ADAM for optimization.

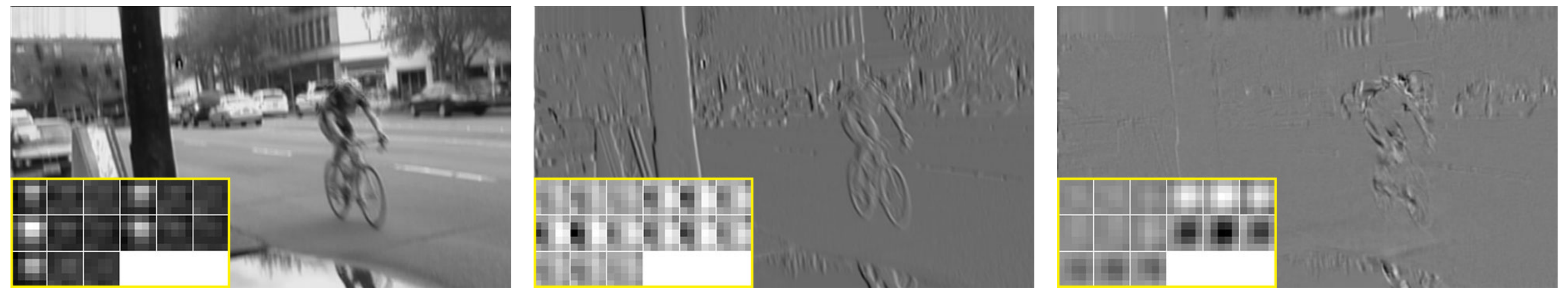
A selection of blurry/sharp pairs (split left/right respectively) from our ground truth dataset.

Effects of Multi-frame: The contribution of using a temporal window is analyzed by keeping the same network architecture, but replicating the central reference frame instead of stacking neighboring frames. The network is retrained with the same hyper-parameters.

Effects of Alignment: Differently learned configurations based on alignment types are analyzed: no-alignment, homography alignment, and optical flow alignment. Quantitative results show the necessity of applying different alignment varying input blurriness.



Analysis on Learned Filters: To gain some insights of what DeblurNet learns, we visualize three representative filters at F0. These are shown to preserve color tone, extract edges, and detect warping artifacts respectively.



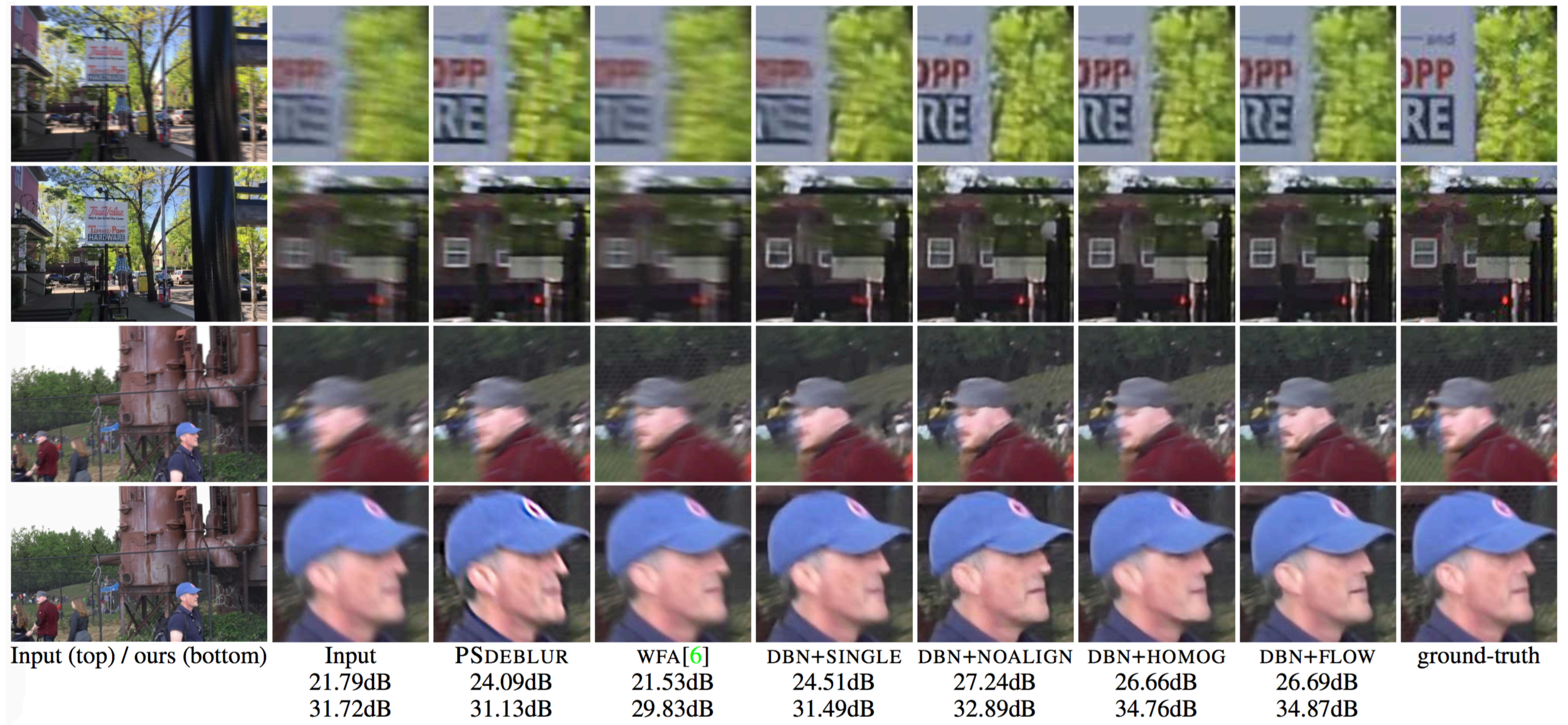
Tone Edge Artifacts

Results

Quantitative Comparison:

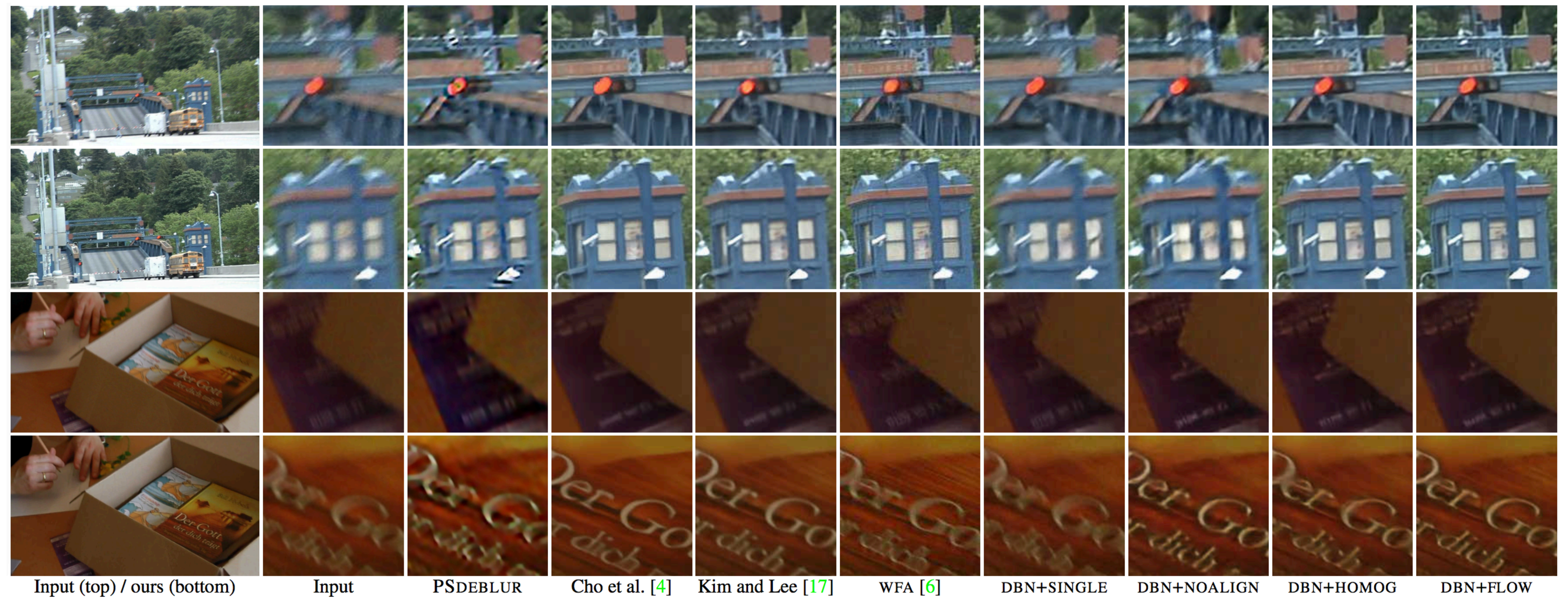
Method	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Average
Input	24.14 / .859	30.52 / .958	28.38 / .914	27.31 / .900	22.60 / .852	29.31 / .951	27.74 / .939	23.86 / .906	30.59 / .976	26.98 / .926	27.14 / .918
PSDEBLUR	24.42 / .908	28.77 / .952	25.15 / .928	27.77 / .928	22.02 / .890	25.74 / .932	26.11 / .948	19.75 / .822	26.48 / .963	24.62 / .938	25.08 / .921
WFA [6]	25.89 / .910	32.33 / .974	28.97 / .931	28.36 / .925	23.99 / .910	31.09 / .975	28.58 / .955	24.78 / .926	31.30 / .981	28.20 / .960	28.35 / .944
DBN+SINGLE	25.75 / .901	31.15 / .966	29.30 / .946	28.38 / .922	23.63 / .885	30.70 / .962	29.23 / .959	25.62 / .936	31.92 / .983	28.06 / .949	28.37 / .941
DBN+NOALIGN	27.83 / .940	**33.11** / .980	**31.29 / .973**	29.73 / .948	25.12 / .930	**32.52** / .978	**30.80 / .975**	**27.28** / .962	**33.32 / .989**	29.51 / .969	**30.05** / .964
DBN+HOMOG.	27.93 / .945	32.39 / .975	30.97 / .969	29.82/ .948	24.79 / .925	31.84 / .972	30.46 / .972	26.64 / .955	33.15 / **.989**	29.30 / .969	29.73 / .962
DBN+FLOW	**28.31 / .956**	33.14 / **.982**	30.92 / **.973**	**29.99 / .954**	**25.58 / .944**	32.39 / **.981**	30.56 / **.975**	27.15 / **.963**	32.95 / **.989**	**29.53 / .975**	**30.05 / .969**

PSNR/MSSIM measurements for each approach, averaged over all frames, for 10 test datasets.



Quantitative results from our test set, with PSNRs relative to the ground truth.

Qualitative Comparison:



Qualitative comparisons to existing approaches.