Online Lens Motion Smoothing for Video Autofocus
Supplemental Materials
Ablation Study

Abdullah Abuolaim[1]
[1]York University, Toronto
abuolaim@eecs.yorku.ca

Michael S. Brown[1,2]
[2]Samsung AI Center, Toronto
mbrown@eecs.yorku.ca

S1. Ablation Study

The supplemental materials provide an ablation study of two variations of our BLSTM architecture: (1) a traditional LSTM architecture with one direction (forward pass only) in Section S2 and (2) our BLSTM with different input sizes in Section S3. This is related to Section 3.3 in the main paper. Based on the main papers notation, MA stands for the moving average *offline* method and SG stands for the Savitzky-Golay *offline* method.

Data Preparation For this ablation study, we first apply both *offline* MA and SG *offline* smoothing for each of the 36 videos. For the MA method, we set the window size $l = 41$ to apply a reasonable smoothing. For the SG method, we set window size $l = 141$ and polynomial degree $m = 3$ in order to impose strong smoothing with some flexibility. Next, we take the output smoothed lens positions to prepare the data for the BLSTM model as described in Section 3.3 of the main paper. We slide a time window of size 20 and stride of 1 to form our BLSTM input samples $\mathbf{X}_i$ and their corresponding y_i for each video data. We finally obtain two types of BLSTM models: (1) BLSTM-MA trained using the data produced by MA and (2) BLSTM-SG trained using the data produced by SG.

S2. Traditional LSTM vs. Bidirectional LSTM

Our first study, uses a traditional LSTM architecture with one direction (forward pass). This is referred to as a unidirectional LSTM (ULSTM). We use the same settings and hyperparameters used for our BLSTM, the only change is that we adopt a forward pass LSTM which results in a 32-dimensional fully connected layer at the top. This study is provided to investigate the effect of having a bidirectional LSTM and its ability to learn from two passes (i.e., forward and backward). In Table 1, we report the lens motion reduction and sharpness change of two *online* LSTM models (i.e., BLSTM and ULSTM) applied for different AF objectives. Overall, BLSTM has reduced lens motion more compared to ULSTM, except for BLSTM-MA on 9 FP and BLSTM-SG on FR. Moreover, BLSTM for all objectives has much smaller loss in sharpness compared to ULSTM. This Table shows that our BLSTM model has an advantage over the ULSTM, and learning from two passes has improved the LSTM's ability to learn smoothing patterns from *offline* smoothing methods with a slight loss in sharpness.

Objective	Lens Motion Reduction			
	Trained Using MA Data		Trained Using SG Data	
	BLSTM	ULSTM	BLSTM	ULSTM
Global	**43.13%**	31.81%	**33.38%**	29.69%
9 FP	24.52%	**26.43%**	**63.61%**	43.83%
51 FP	**36.66%**	34.75%	**40.75%**	37.23%
FR	**25.28%**	20.85%	11.20%	**18.75%**

Objective	Sharpness Change			
	Trained Using MA Data		Trained Using SG Data	
	BLSTM	ULSTM	BLSTM	ULSTM
Global	**-0.29%**	-0.39%	**-0.47%**	-0.84%
9 FP	**-1.15%**	-1.97%	**-2.25%**	-2.42%
51 FP	**-4.04%**	-5.69%	**-5.17%**	-6.14%
FR	**-0.45%**	-1.17%	**-1.55%**	-7.55%

Table 1: A comparison between different LSTM architectures: bidirectional (BLSTM) vs. unidirectional (ULSTM). This table shows lens motion reduction and its effect on sharpness after applying different *online* LSTM smoothing methods. Compared to ULSTM, BLSTM has a better ability to learn smoothing patterns from *offline* smoothing methods with only a slight drop in sharpness.

S3. BLSTM with Different Input Sizes

Next, we examine a variant of our proposed method that uses the same architecture of our BLSTM, but adjusts the input size of time window l. We introduce our BLSTM with

Objective	Lens Motion Reduction					
	Trained Using MA Data			Trained Using SG Data		
	$BLSTM_{10}$	$BLSTM_{20}$	$BLSTM_{40}$	$BLSTM_{10}$	$BLSTM_{20}$	$BLSTM_{40}$
Global	**50.31%**	43.13%	28.90%	**43.45%**	33.38%	31.05%
9 FP	**44.02%**	24.52%	30.51%	**66.84%**	63.61%	16.47%
51 FP	**41.21%**	36.66%	29.04%	**58.96%**	40.75%	33.65%
FR	**31.63%**	25.28%	22.42%	06.33%	11.20%	**19.96%**

Objective	Lens Motion Reduction					
	Trained Using MA Data			Trained Using SG Data		
	$BLSTM_{10}$	$BLSTM_{20}$	$BLSTM_{40}$	$BLSTM_{10}$	$BLSTM_{20}$	$BLSTM_{40}$
Global	-0.43%	**-0.29%**	-0.30%	-1.06%	**-0.47%**	-0.61%
9 FP	-2.67%	-1.15%	**-0.58%**	-3.05%	**-2.25%**	-3.55%
51 FP	-4.52%	**-4.04%**	-4.25%	-8.76%	**-5.17%**	-5.66%
FR	-0.71%	**-0.45%**	-0.87%	-3.15%	-1.55%	**-1.00%**

Table 2: A comparison of our BLSTM with three different input sizes: $BLSTM_{10}$ (l =10), $BLSTM_{20}$ (l =20), and $BLSTM_{40}$ (l =40). This table shows lens motion reduction and its effect on sharpness after applying different *online* BLSTM smoothing methods. $BLSTM_{20}$ imposes a reasonable reduction in lens motion and at the same time has the smallest loss in sharpness for the most cases.

three different input sizes: $BLSTM_{10}$ (l =10), $BLSTM_{20}$ (l =20 as we set in the main paper), and $BLSTM_{40}$ (l =40). Table 2 presents the results of applying BLSTM with vary input size and shows the amount of lens motion reduction with its effect on sharpness. By looking at the lens motion reduction table, BLSTM with a smaller window size imposes a larger reduction of lens motion in general. However, BLSTM with the smallest window size (i.e., $BLSTM_{10}$) also introduces more loss in sharpness compared to others. $BLSTM_{20}$ imposes a reasonable reduction in lens motion and at the same time has the smallest loss in sharpness for the most cases.

S4. Summary

Our ablation study results show that our proposed *online* BLSTM is able to learn smoothed lens motion patterns from different *offline* methods (i.e., MA and SG) and performs almost as well as *offline* methods in an *online* manner.