Online Lens Motion Smoothing for Video Autofocus Supplemental Materials Results

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

Our supplemental materials provide additional results including: BLSTM accuracy of each objective for different scenes (Section S1), qualitative results of both *offline* SG and *online* BLSTM applied on all the 36 videos (Section S2), and WMA qualitative results for 36 videos (Section S3).

S1. BLSTM Accuracy

As described in Section 4 of the main paper, we divided the scenes into pairs, then performed cross-validation where we take out-of-sample a pair of scenes (one for validation and another for testing). Under this regime, we have trained a total of 21 models.

Accuracy is measured as how similar our BLSTM predication is to the *offline* smoothed lens motion. Table 1 shows BLSTM accuracy of each objective for different scenes. Our accuracy is reasonable, and is especially good for the global objective. This is because the global objective has a single region of interest (ROI) and usually involves less number of lens movements. Compared to other objectives, 51-FP has slightly lower scene accuracy due to the fact it has 51 ROIs and usually involves more lens movements.

In general, these results show that our proposed BLSTM is able to learn smoothed lens motion patterns and able to generalize for different scenes by testing on independent scenes that have been never seen by the network during training.

Objective	Model	Scene	Accuracy	Objective	Model	Scene	Accuracy
Global	$\mathbf{G}_{1,8}$	1	0.892	9-FP	$\mathbf{9\text{-}FP}_{1,8}$	1	0.982
		8	0.996			8	0.879
	$\mathbf{G}_{2,7}$	2	0.961		$\mathbf{9\text{-}FP}_{2,7}$	2	0.827
		7	0.994			7	0.941
	$\mathbf{G}_{3,6}$	3	0.986		$\mathbf{9\text{-}FP}_{3,6}$	3	0.875
		6	0.988			6	0.954
	$\mathbf{G}_{4,10}$	4	0.994		$\mathbf{9\text{-}FP}_{4,10}$	4	0.996
		10	0.989			10	0.899
	$\mathbf{G}_{5,9}$	5	0.993		$\mathbf{9\text{-}FP}_{5,9}$	5	0.986
		9	0.992			9	0.884
FR	–	1	–	51-FP	$\mathbf{51\text{-}FP}_{1,8}$	1	0.716
	–	2	–			8	0.888
	$\mathbf{FR}_3$	3	0.992		$\mathbf{51\text{-}FP}_{2,7}$	2	0.824
	$\mathbf{FR}_4$	4	0.810			7	0.959
	–	5	–		$\mathbf{51\text{-}FP}_{3,6}$	3	0.924
	$\mathbf{FR}_6$	6	0.870			6	0.865
	$\mathbf{FR}_7$	7	0.984		$\mathbf{51\text{-}FP}_{4,10}$	4	0.773
	$\mathbf{FR}_8$	8	0.986			10	0.825
	–	9	–		$\mathbf{51\text{-}FP}_{5,9}$	5	0.778
	$\mathbf{FR}_{10}$	10	0.943			9	0.816

Table 1: BLSTM accuracy of each objective for different scenes. Overall, scene accuracy for all is mostly high especially for the global objective, because it always has a single ROI and usually involves less number of lens movements. 51-FP has slightly lower scene accuracy, because it has 51 ROIs and usually involves more lens movements. Recall, FR objective can be applied on scenes that have faces and only 6 scenes contain faces.

S2. BLSTM Online Performance

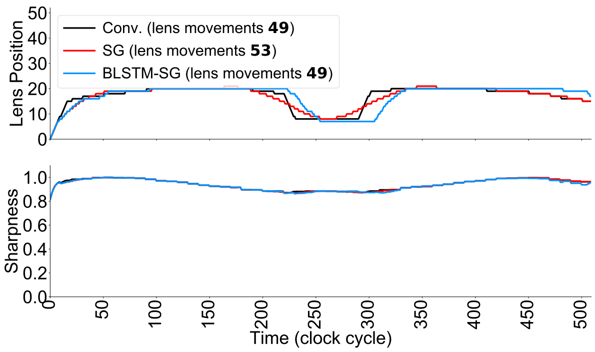
BLSTM accuracy described above evaluates the model based on the *offline* ground truth data. This accuracy metric suffers from cumulative error for *online* evaluations. For example, constructing lens movement trajectory for the whole video using those *offline* evaluated samples results in propagating the error of any misclassified sample. To that end, we adopt the trained BLSTM model to predict *online* stream data of consecutive lens positions. This *online* BLSTM receives the optimal value o_t coming from phase difference module at each time step t where the lens position x_t is predicted based on the previously observed sample $\mathbf{X}_{t-1}$. In a repeated way and by receiving o_t at a time, our *online* BLSTM is able to build the lens movements for the whole video.

In Fig. 1 and Fig. 2, we present the results of both *offline* SG and *online* BLSTM applied on all the 36 videos as described in Section 4 of the main paper. These figures show the lens positions and the corresponding sharpness values over time. Total number of lens movements for each method is shown in the plot's legend between parentheses.

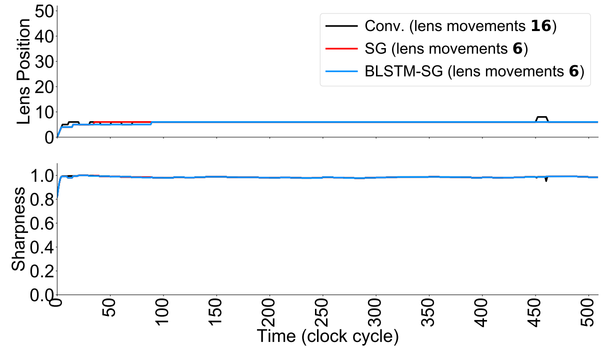
These qualitative results show that our *online* BLSTM produces lens motion trajectories very similar to the *offline* method trajectories used for training. In addition, our *online* BLSTM has significantly fewer lens movements compared to conventional AF algorithms.

S3. WMA Online Performance

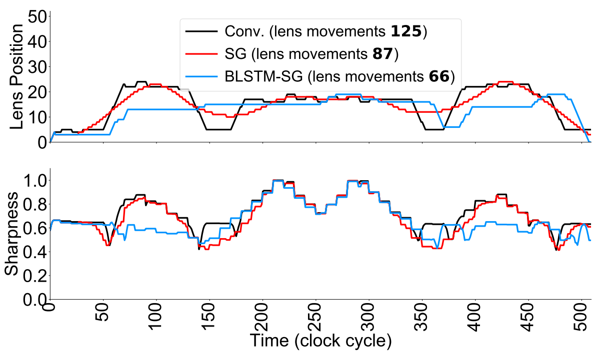
In section 4 of the main paper, we only presented WMA qualitative results of one video. The remaining results of the *online* WMA for the 36 videos are presented in Fig. 3 and Fig. 4. These figures show the lens positions and the corresponding sharpness values over time. The *online* WMA is able to suppress the small lens fluctuations of the conventional lens positions while maintaining reasonable sharpness values. The WMA shows a strong ability in preserving the sharpness of the ROI, because the WMA was designed in a way to weight lens position by phase difference optimal value that has the lens position of the maximum sharpness.



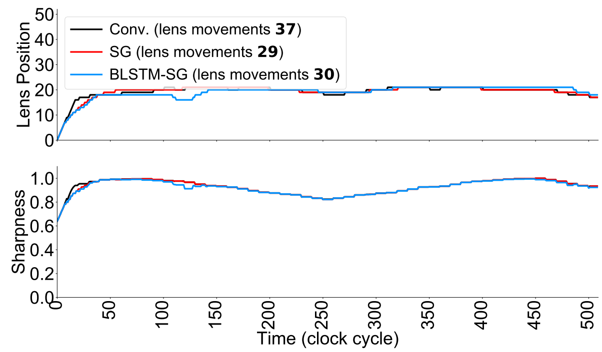
(a) Scene 1, global objective



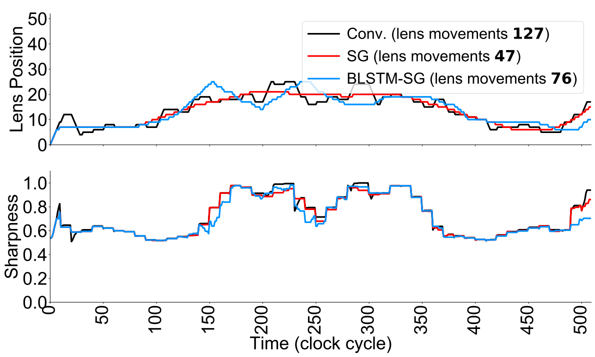
(b) Scene 1, 9-FP objective



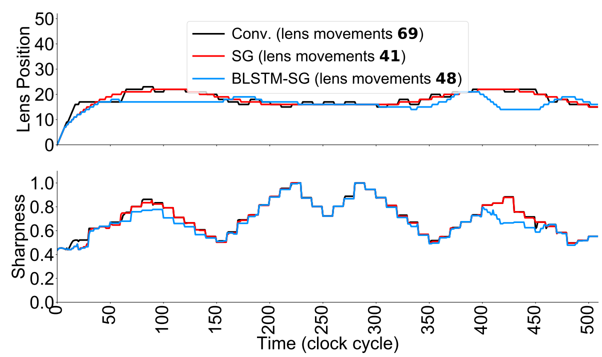
(c) Scene 1, 51-FP objective



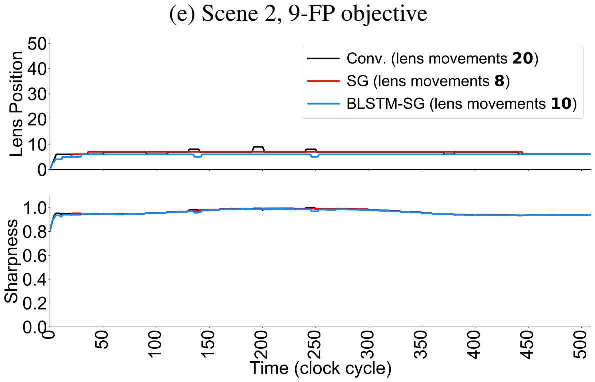
(d) Scene 2, global objective



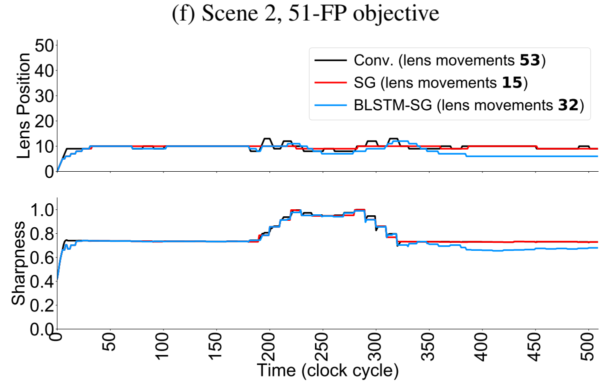
(e) Scene 2, 9-FP objective



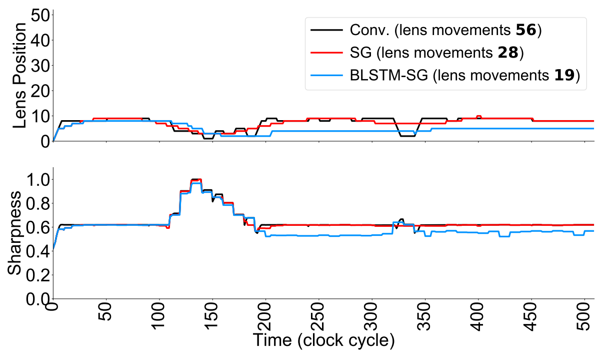
(f) Scene 2, 51-FP objective



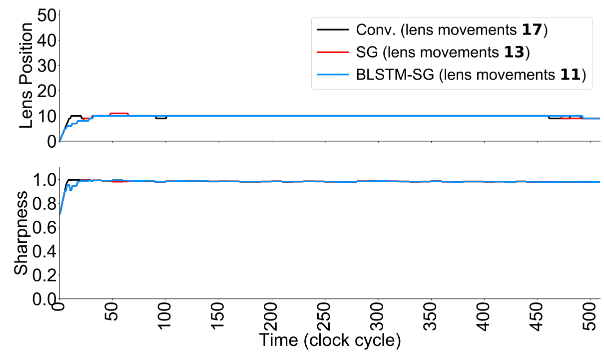
(g) Scene 3, global objective



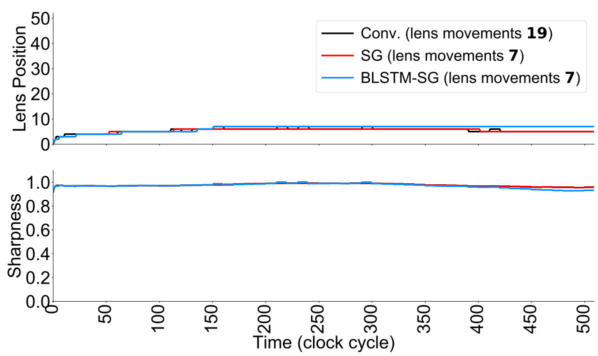
(h) Scene 3, 9-FP objective



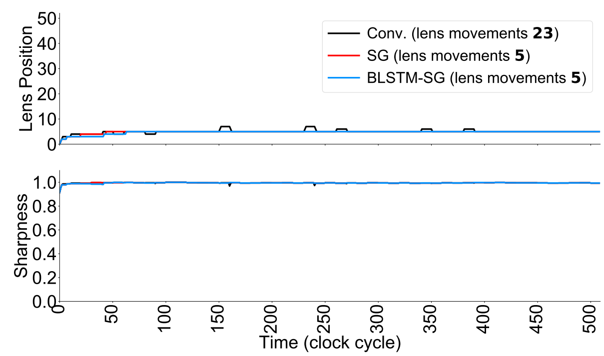
(i) Scene 3, 51-FP objective



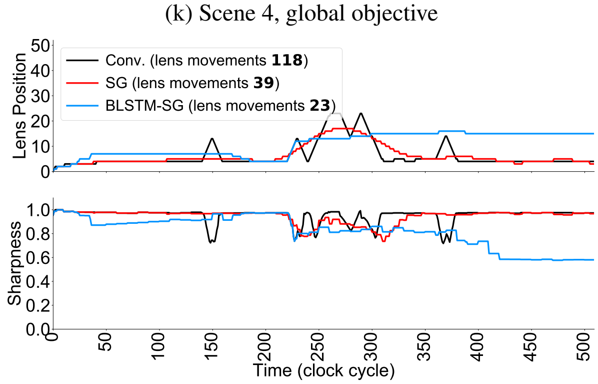
(j) Scene 3, FD objective



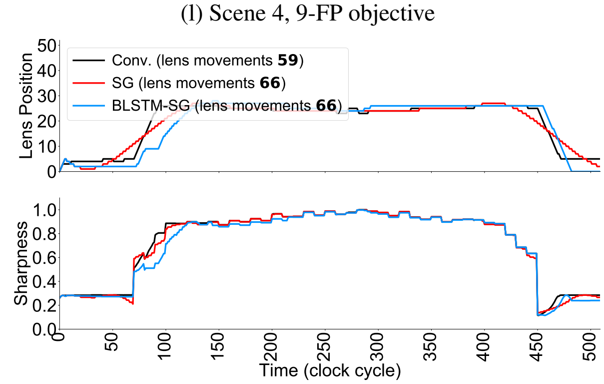
(k) Scene 4, global objective



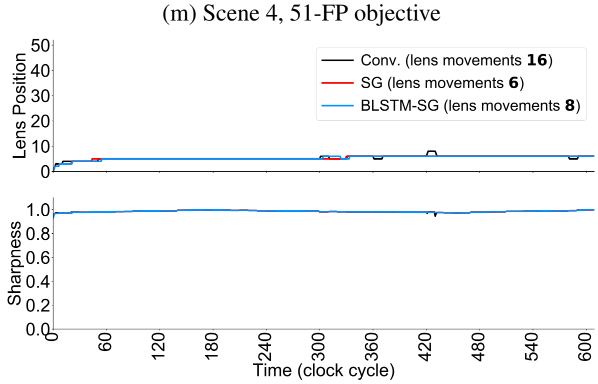
(l) Scene 4, 9-FP objective



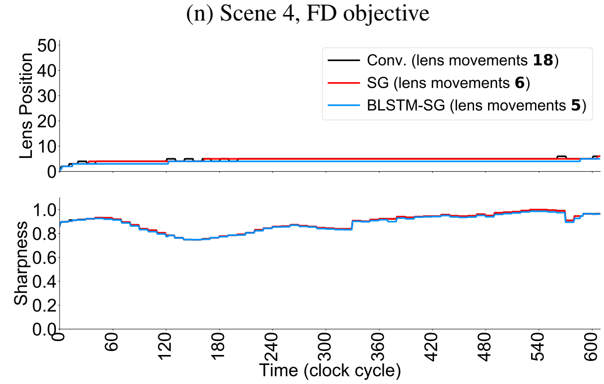
(m) Scene 4, 51-FP objective



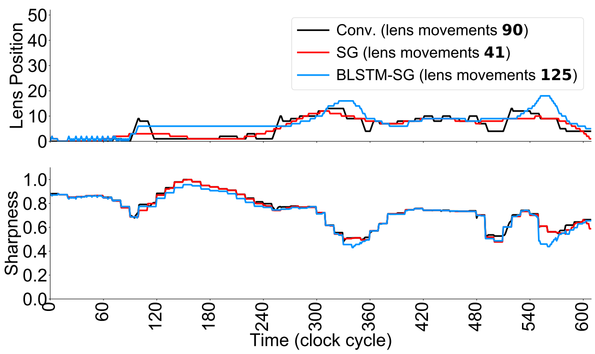
(n) Scene 4, FD objective



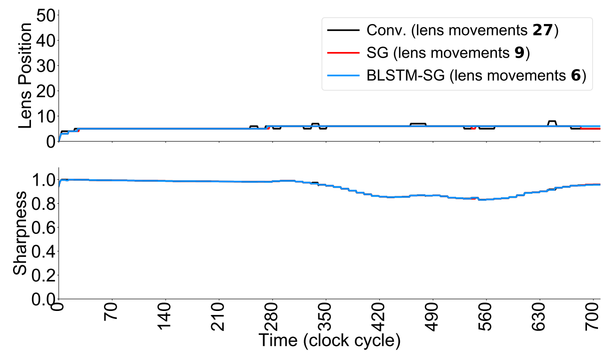
(o) Scene 5, global objective



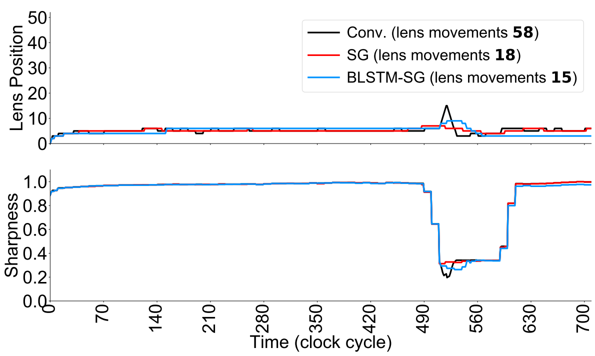
(p) Scene 5, 9-FP objective



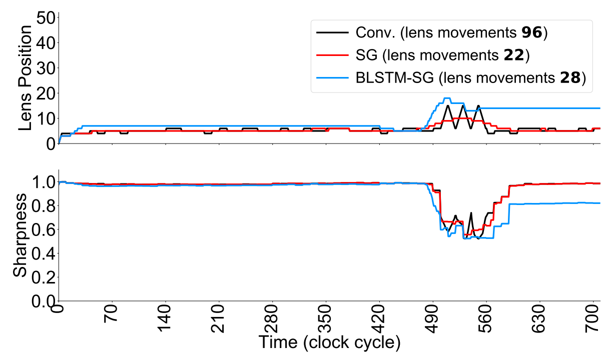
(q) Scene 5, 51-FP objective



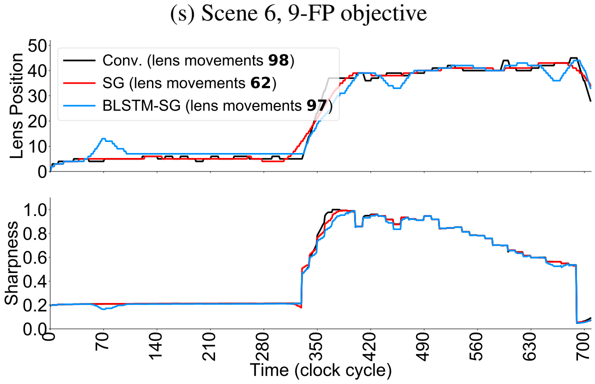
(r) Scene 6, global objective



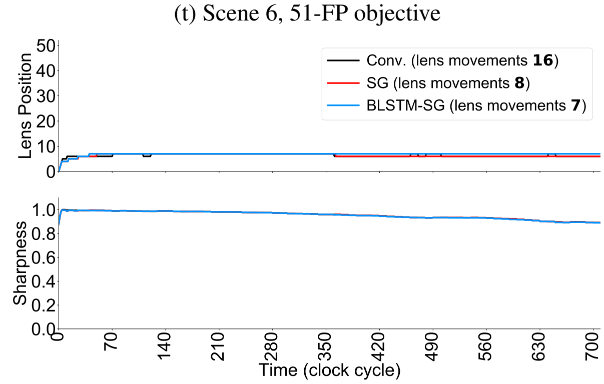
(s) Scene 6, 9-FP objective



(t) Scene 6, 51-FP objective

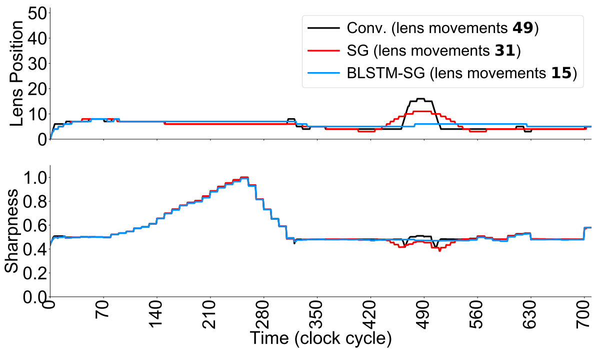


(u) Scene 6, FD objective

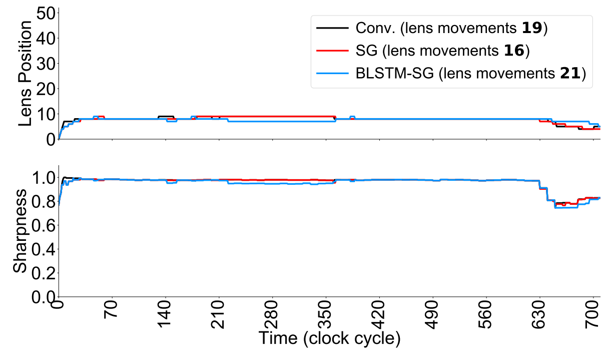


(v) Scene 7, global objective

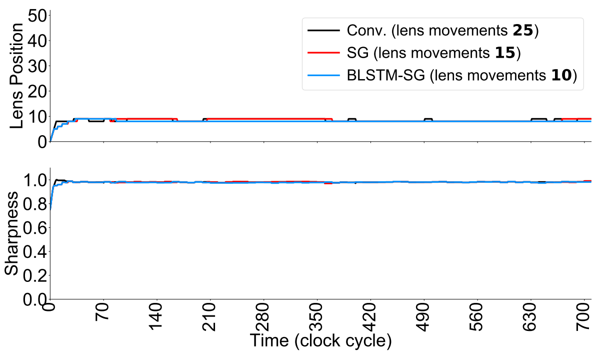
Figure 1: A comparison of lens positions between the conventional and BLSTM. SF is also shown as this method is used to train the BLSTM model. The plot above presents the lens positions over time and the one below shows the effect on sharpness value on the same timeline. Total number of lens movements for each method is shown in the plot's legend between parentheses. BLSTM is able to suppress small lens fluctuations without substantially affecting the sharpness.



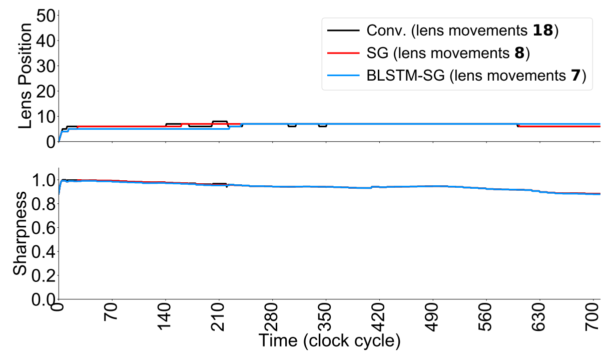
(a) Scene 7, 9-FP objective



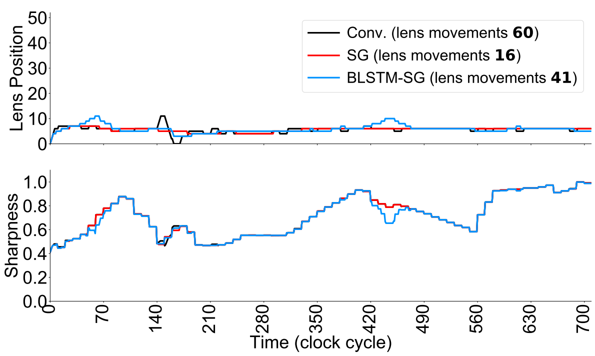
(b) Scene 7, 51-FP objective



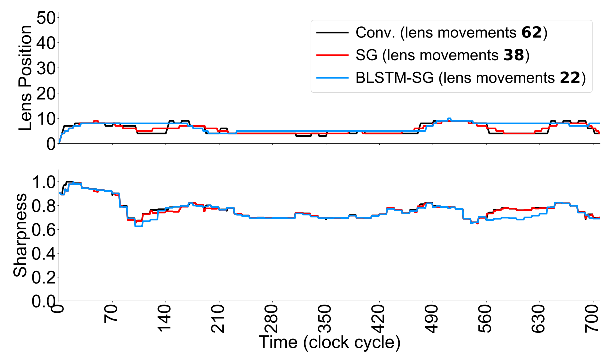
(c) Scene 7, FD objective



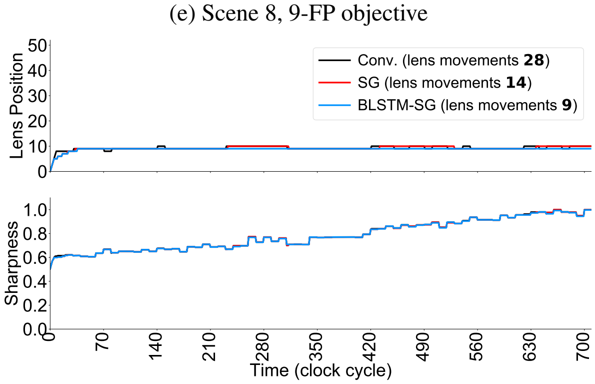
(d) Scene 8, global objective



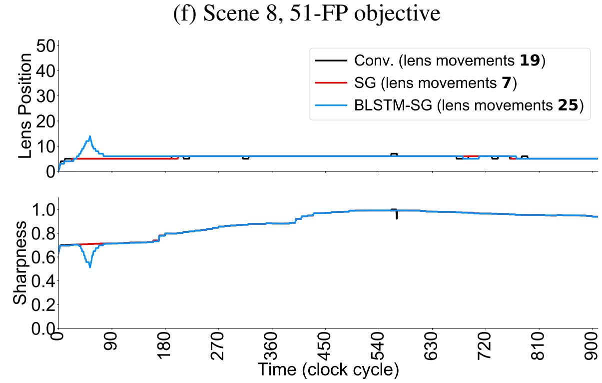
(e) Scene 8, 9-FP objective



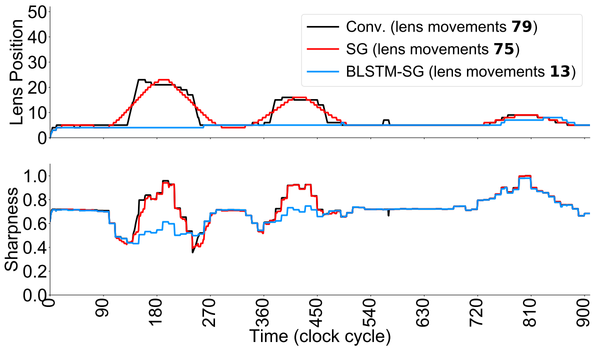
(f) Scene 8, 51-FP objective



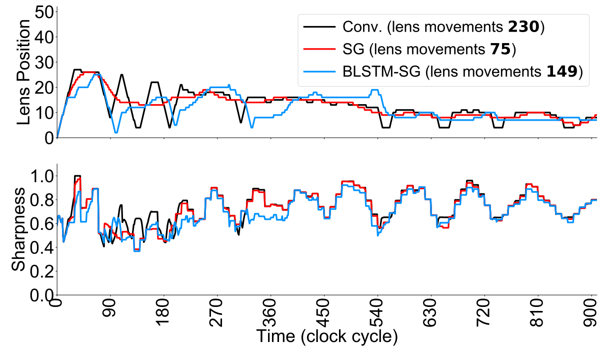
(g) Scene 8, FD objective



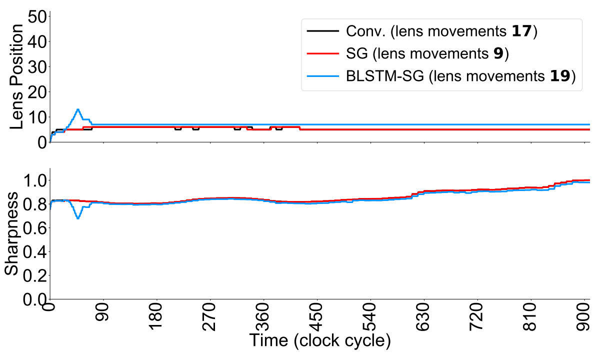
(h) Scene 9, global objective



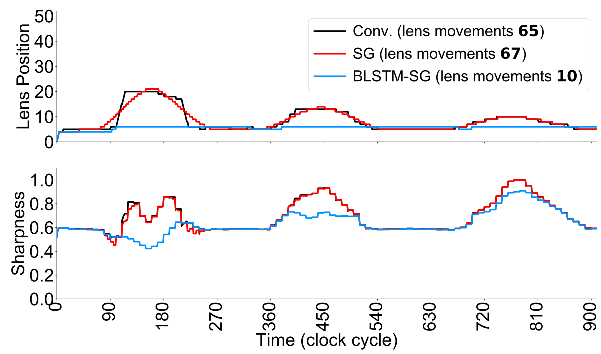
(i) Scene 9, 9-FP objective



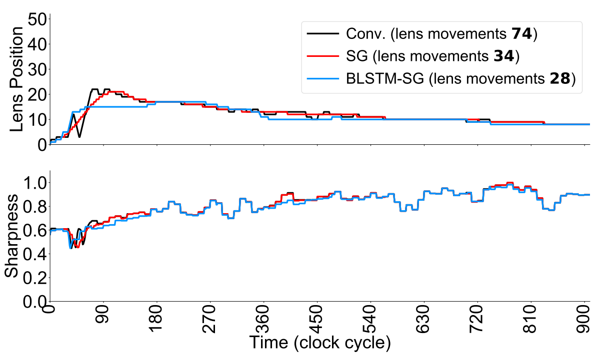
(j) Scene 9, 51-FP objective



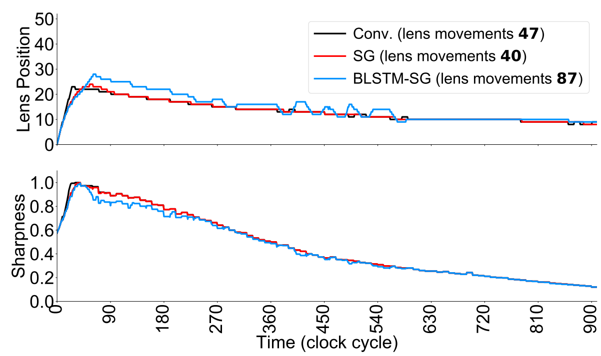
(k) Scene 10, global objective



(l) Scene 10, 9-FP objective

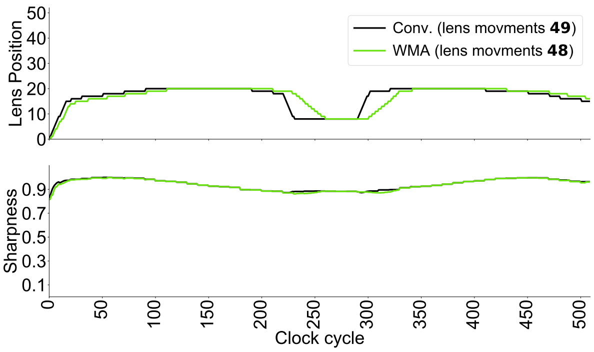


(m) Scene 10, 51-FP objective

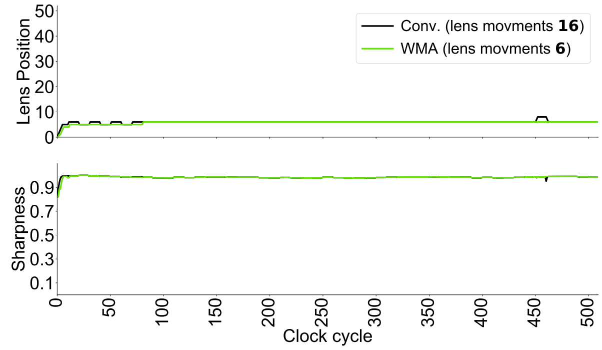


(n) Scene 10, FD objective

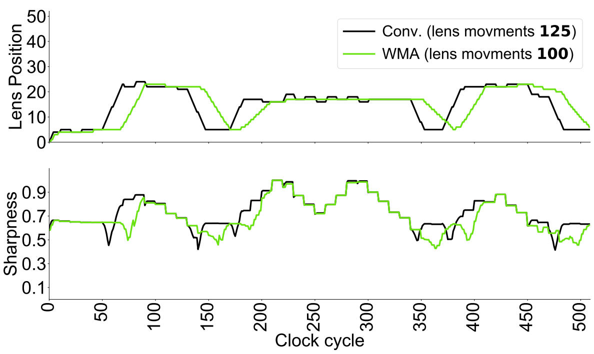
Figure 2: A comparison of lens positions between the conventional and BLSTM. SF is also shown as this method is used to train the BLSTM model. The plot above presents the lens positions over time and the one below shows the effect on sharpness value on the same timeline. Total number of lens movements for each method is shown in the plot's legend between parentheses. BLSTM is able to suppress small lens fluctuations without substantially affecting the sharpness.



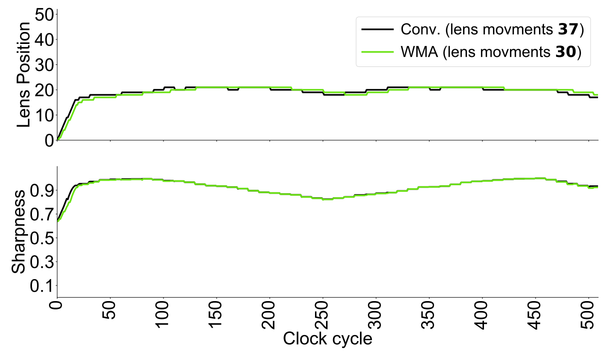
(a) Scene 1, global objective



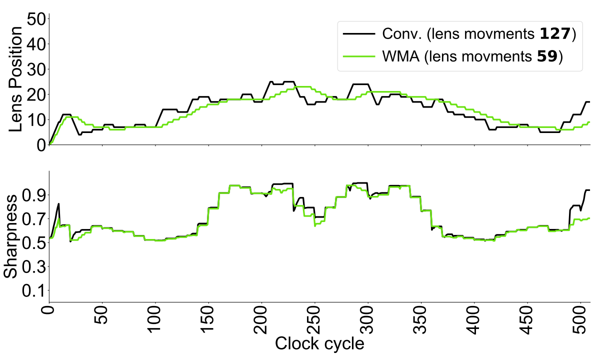
(b) Scene 1, 9-FP objective



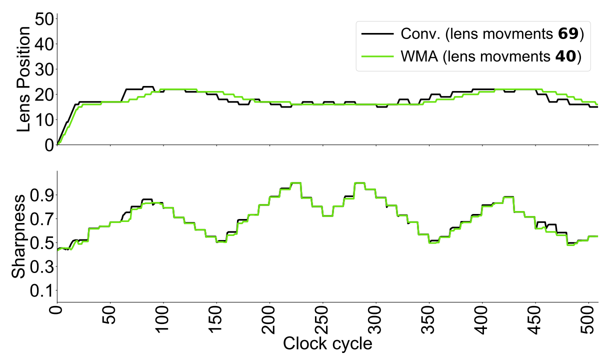
(c) Scene 1, 51-FP objective



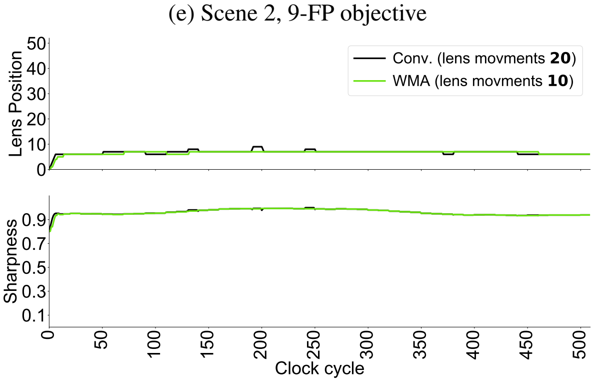
(d) Scene 2, global objective



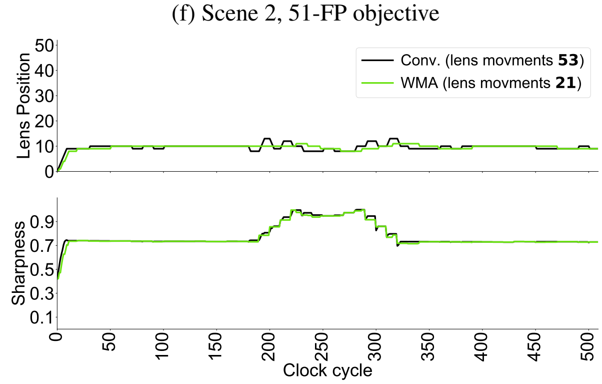
(e) Scene 2, 9-FP objective



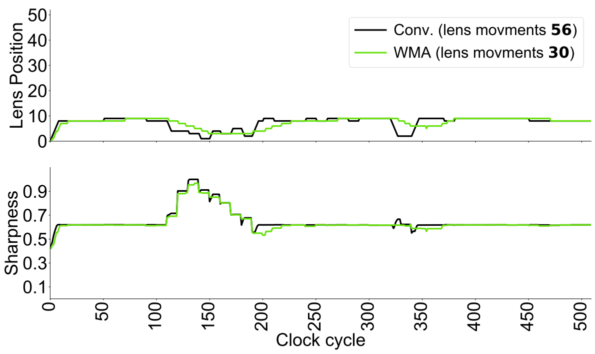
(f) Scene 2, 51-FP objective



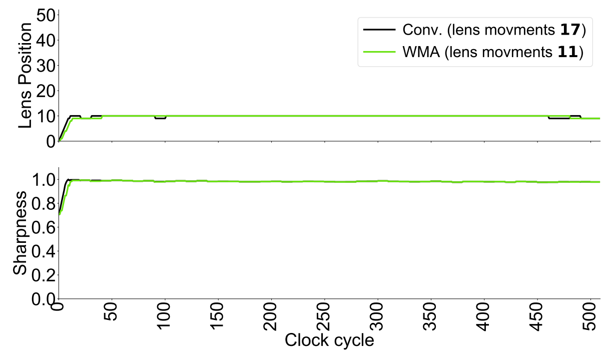
(g) Scene 3, global objective



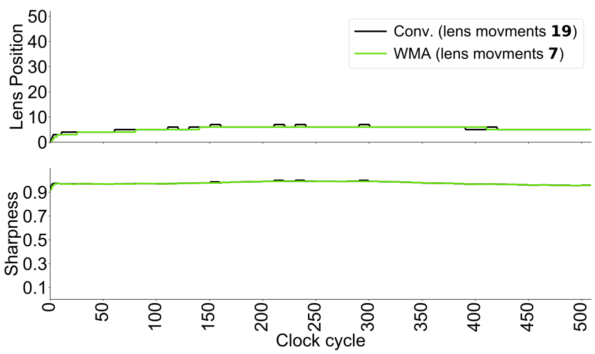
(h) Scene 3, 9-FP objective



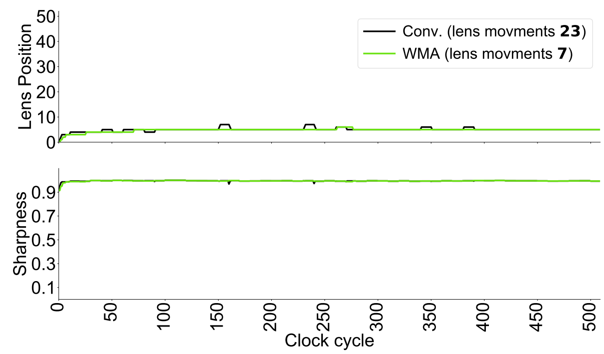
(i) Scene 3, 51-FP objective



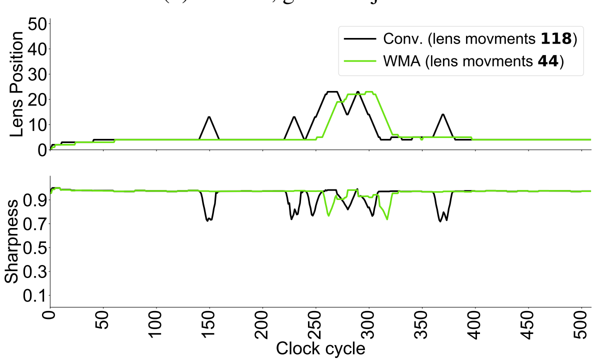
(j) Scene 3, FD objective



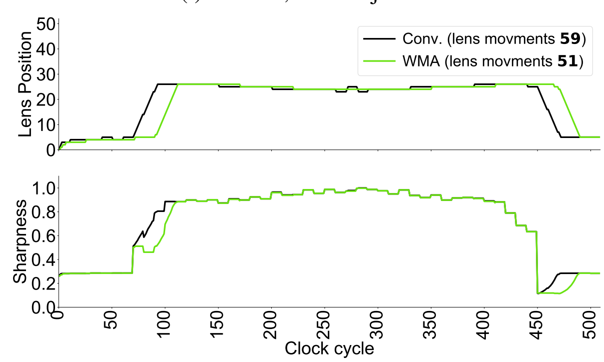
(k) Scene 4, global objective



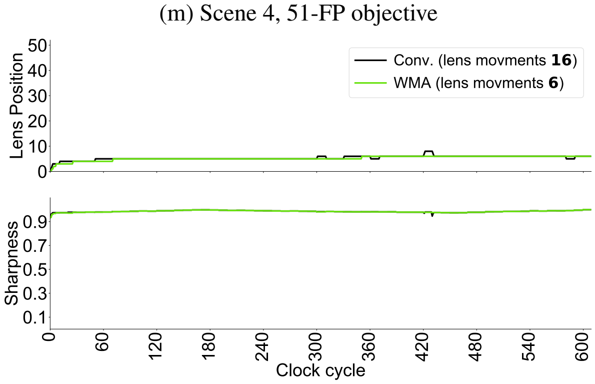
(l) Scene 4, 9-FP objective



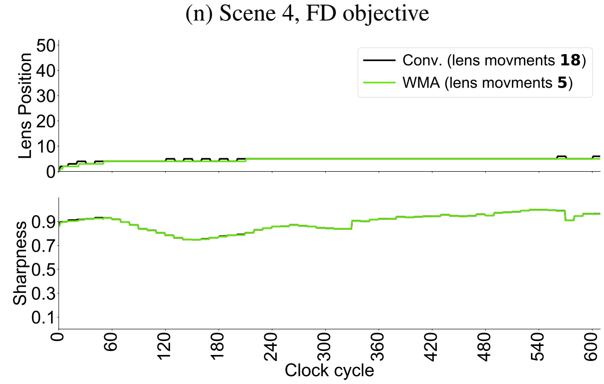
(m) Scene 4, 51-FP objective



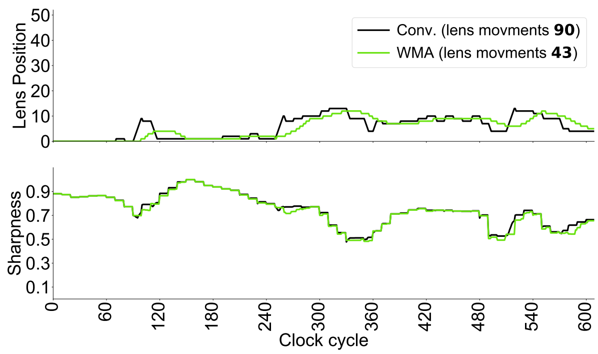
(n) Scene 4, FD objective



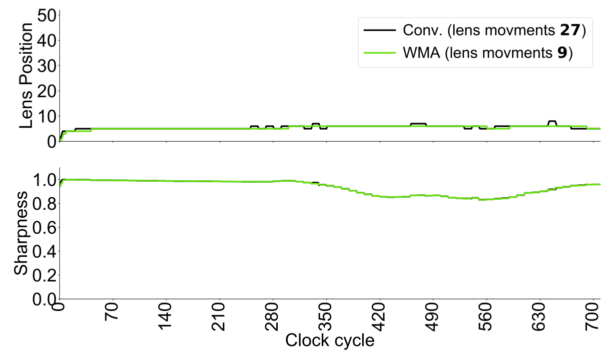
(o) Scene 5, global objective



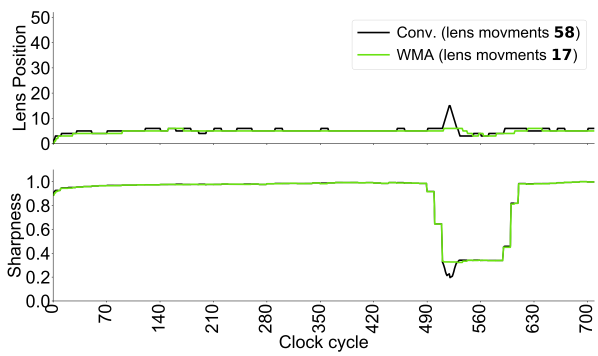
(p) Scene 5, 9-FP objective



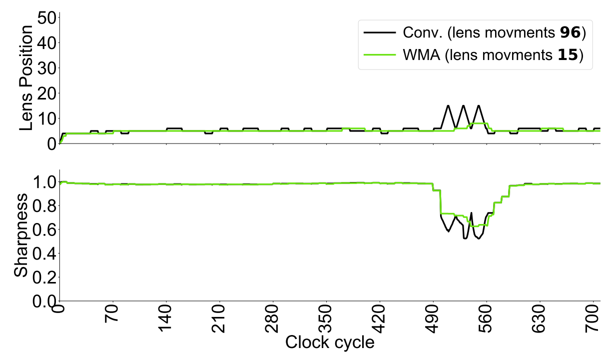
(q) Scene 5, 51-FP objective



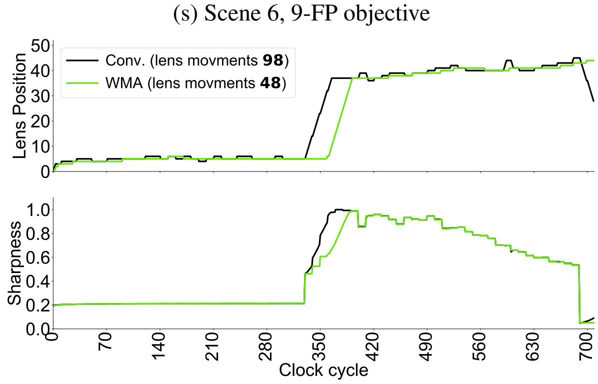
(r) Scene 6, global objective



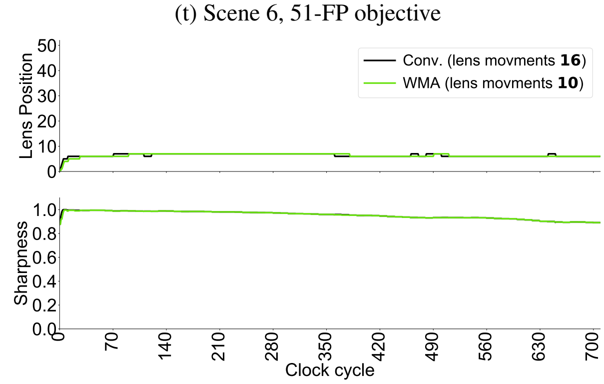
(s) Scene 6, 9-FP objective



(t) Scene 6, 51-FP objective

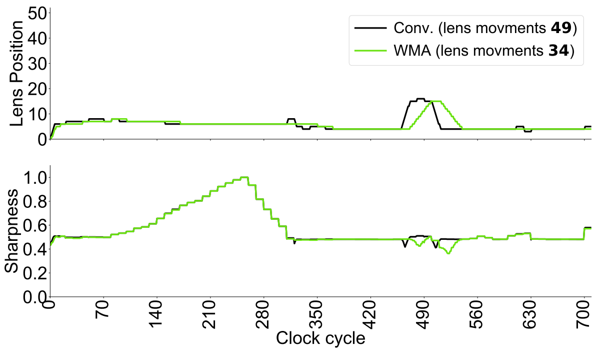


(u) Scene 6, FD objective

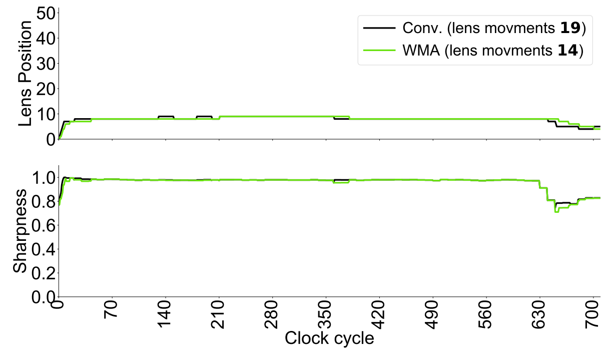


(v) Scene 7, global objective

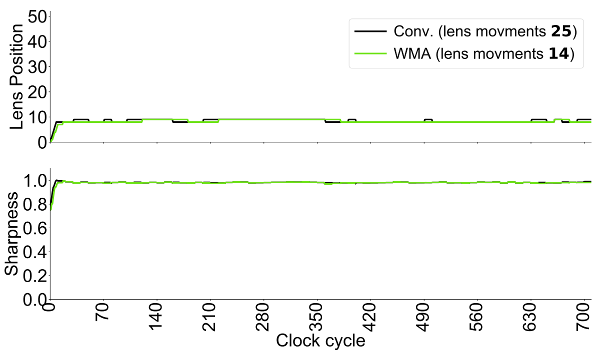
Figure 3: A comparison of lens positions between the conventional and WMA. The plot above presents the lens positions over time and the one below shows the effect on sharpness value on the same timeline. Total number of lens movements for each method is shown in the plot's legend between parentheses. WMA is able to suppress small lens fluctuations without substantially affecting the sharpness.



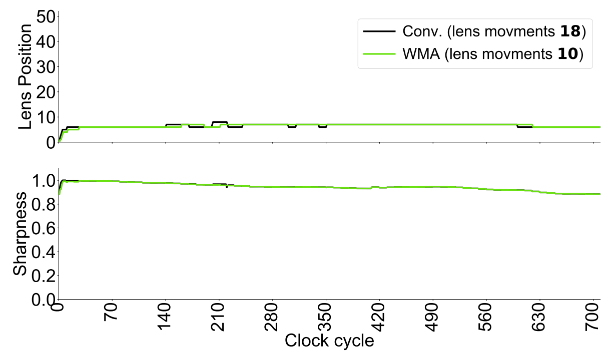
(a) Scene 7, 9-FP objective



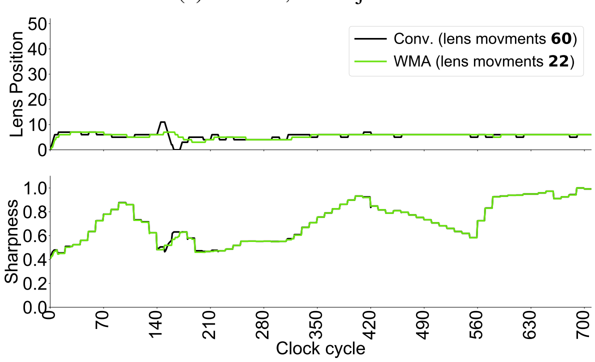
(b) Scene 7, 51-FP objective



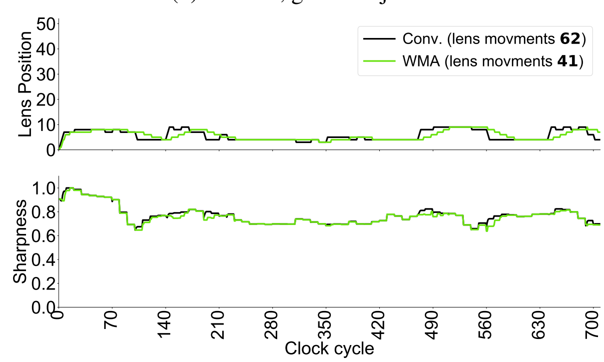
(c) Scene 7, FD objective



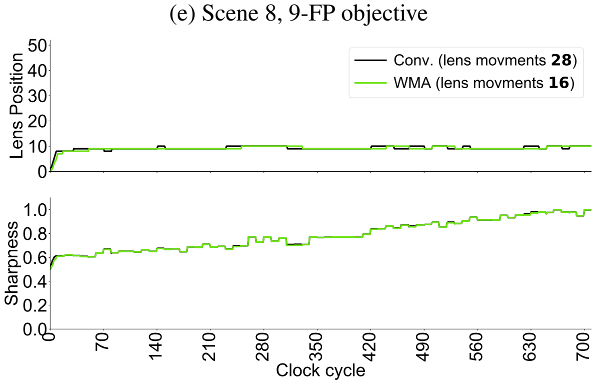
(d) Scene 8, global objective



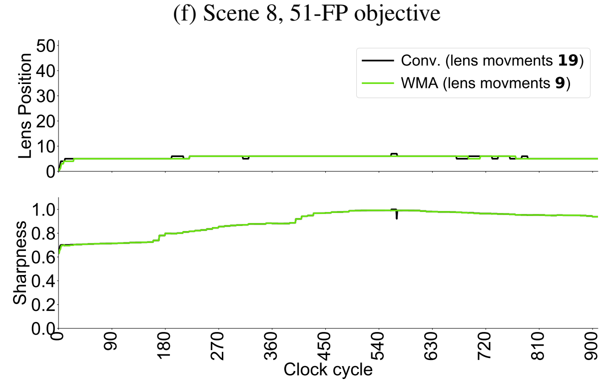
(e) Scene 8, 9-FP objective



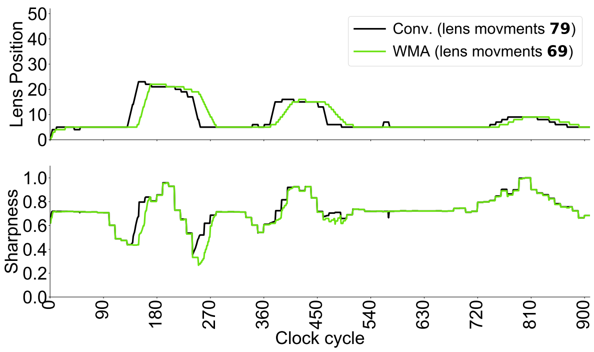
(f) Scene 8, 51-FP objective



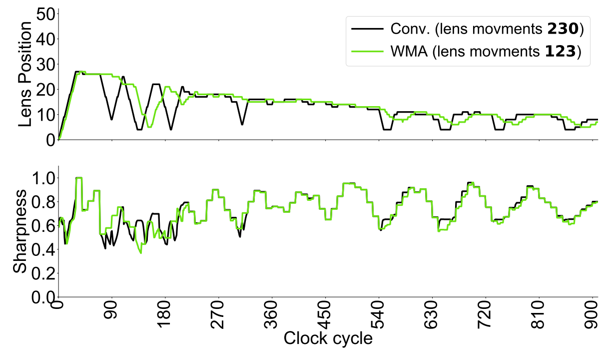
(g) Scene 8, FD objective



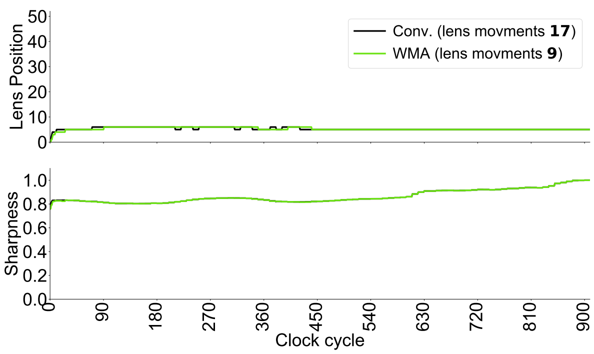
(h) Scene 9, global objective



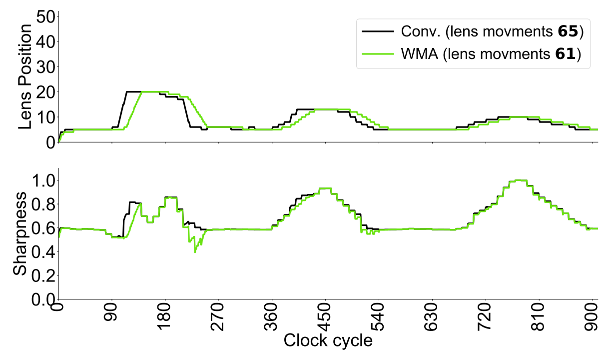
(i) Scene 9, 9-FP objective



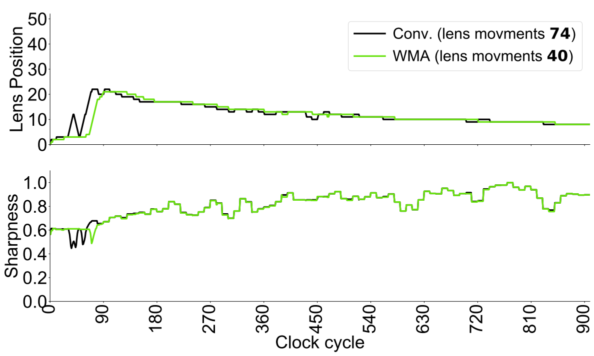
(j) Scene 9, 51-FP objective



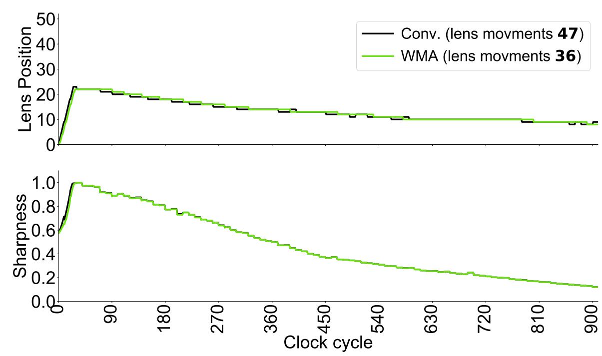
(k) Scene 10, global objective



(l) Scene 10, 9-FP objective



(m) Scene 10, 51-FP objective



(n) Scene 10, FD objective

Figure 4: A comparison of lens positions between the conventional and WMA. The plot above presents the lens positions over time and the one below shows the effect on sharpness value on the same timeline. Total number of lens movements for each method is shown in the plot's legend between parentheses. WMA is able to suppress small lens fluctuations without substantially affecting the sharpness.