Appendix

I. Data collection

The helmet we chose to mount the Azure Kinect RGB-D camera to is a bicycle helmet suitable for head circumferences between 52 and 56 cm, as shown in Fig. I in the supplementary. The helmet itself has adjustable straps and an inner foam shell for comfort. No parts of the helmet appeared in the wearer's field of view.

We limited data collection to 100 actions per recording (approximately 4 minutes to complete) to reduce the possibility of fatigue from wearing the helmet (405 grams) with the Azure Kinect RGB-D camera (440 grams).

Additionally, all cables used during recordings were long enough so that they would not restrict the data collector's movements. The cables were also positioned in a way that would not reduce visibility during recordings: The USB cable for supplying power to the camera was plugged into a 15-feet long extension cord behind the data collector, and the laptop used for recording data was placed as close as possible to the data collector while not obstructing the data collector's view of the rest of the scene. The USB-C data cable connecting the camera to the laptop was also positioned behind the data collector for the same reason as the power cable.

Thanks to the above setup, it does not seem that the volunteers' movement patterns have any noticeable change according to our observation.

II. Visualizations of hand pose estimation

We conducted a quality check for hand pose estimation results, and only the frames with reliable estimations were included in our dataset. Google's MediaPipe Hands tool was quite robust, as most frames had received precise estimations. Some example frames are presented in Fig. II.

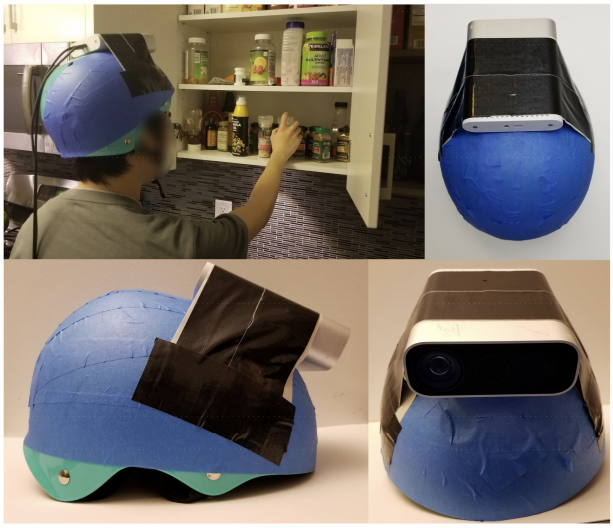

Figure I. Photos of helmets.

Figure II. Visualizations of hand pose estimations.

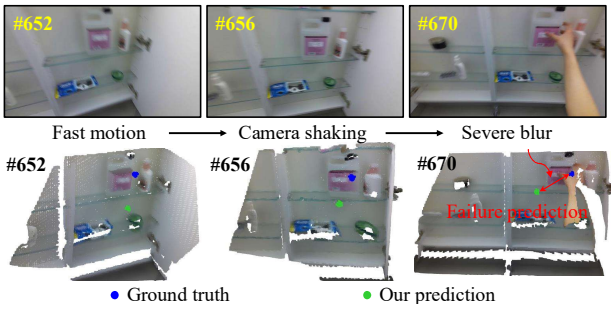

Figure III. Visualization of one failure case.

III. Failure cases

We will add failure case discussions in the final version. Here we show one challenging case with motion blur due to severe viewpoint change in Fig. III. In addition, we will clarify the discussion on motion features and provide analyses about our experiment results.

IV. Visualizations of evaluation results

The visualizations of errors in different time stages are shown in Fig. IV.

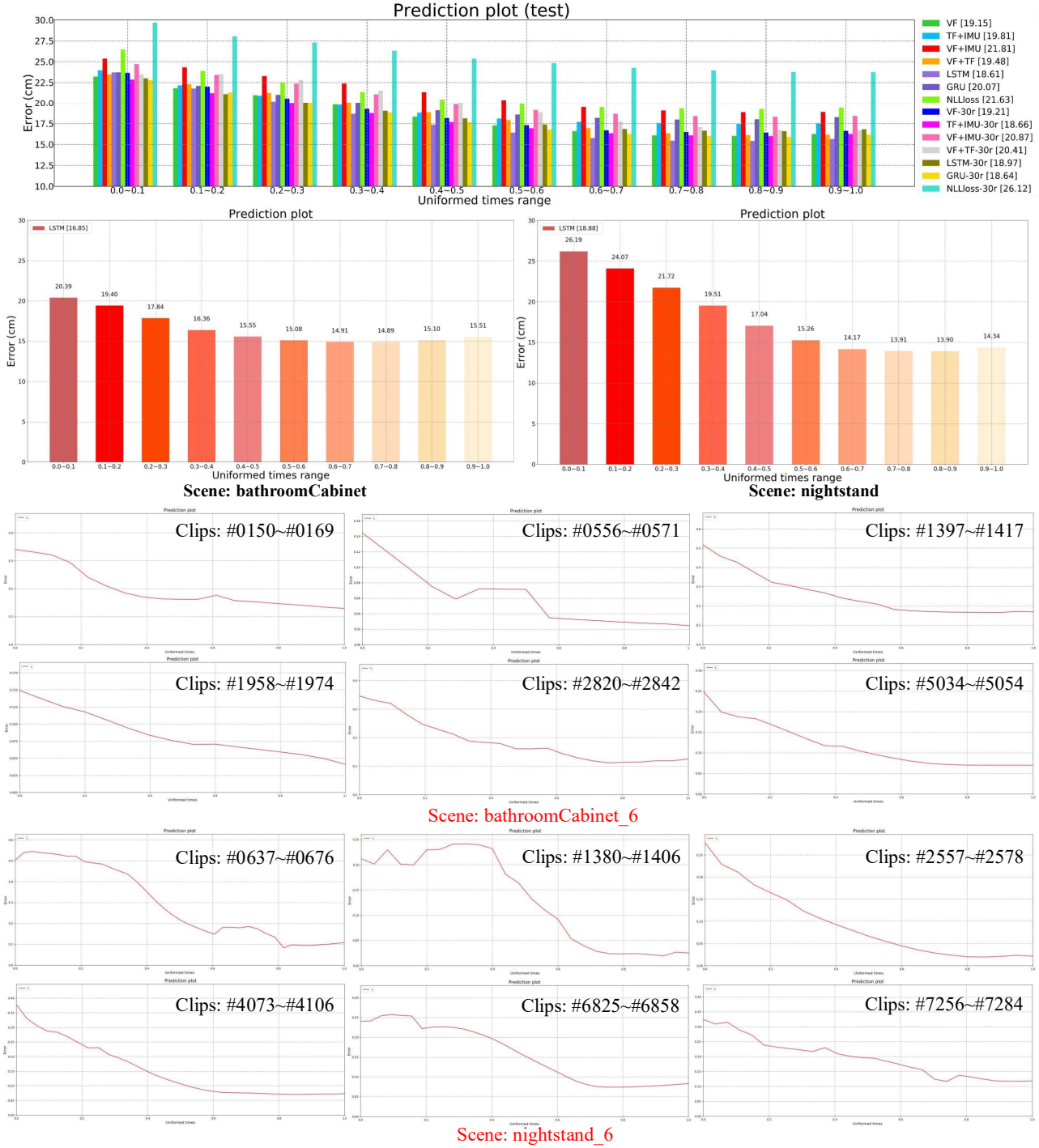

Figure IV. Visualizations of errors in different stages.