	HumanEva [36]	**DIML** [26]	**Evaluation of RGB-D SLAM Systems** [37]	**UTD-MHAD** [9]	**Human Activity Recognition** [2]	**I-MOVE (Ours)**
Velocity Ground Truth						✓
Moving Object(s)	✓				✓	✓
Any Specific Motion Parameter Ground Truth	✓					✓
Images / Video Provided	✓	✓	✓	✓		✓
Indoor Images	✓	✓	✓	✓		✓
Outdoor Images						✓
Multiple Cameras Used	✓					✓
Multiple Perspectives of Same Object	✓					✓
RGB-D		✓	✓			✓
Passive STEREO						✓

Table 1: COMPARISON OF THE CURRENTLY AVAILABLE DATASETS *Above we summarize some of the currently available datasets. It should be observed that none of the available datasets provide a velocity ground truth which is a big contribution of our I-MOVE dataset. The closest dataset to ours is the HumanEva dataset which only contains humans unlike ours which contains a variety of objects. In addition our dataset also contains videos from RGB-D and Passive STEREO cameras.*

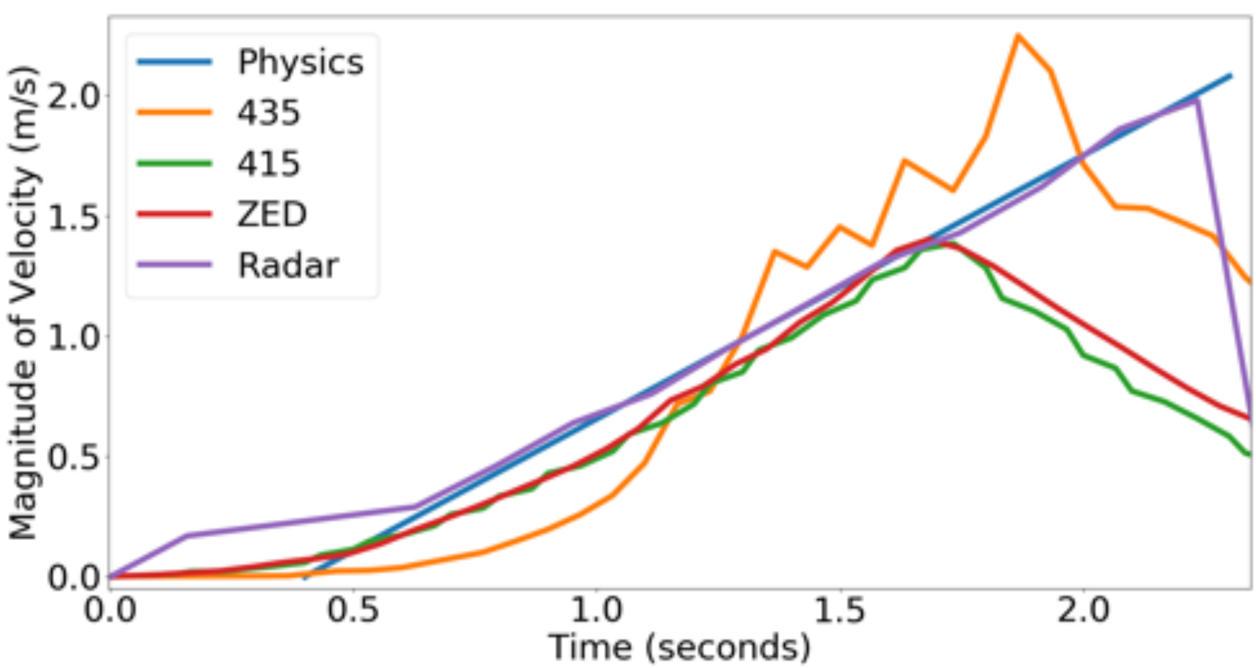

Figure 1: OUTDOOR RAMP SETUP - VELOCITY ESTIMATION *I-MOVE contains many outdoor scenes within the dataset. We conducted some of these outdoor scenes with a similar setup / design to the indoor scenes. One of these replicated scenes includes the rolling object down outdoor ramp / inclined plane. Similar outdoor and indoor scenes such as this provide an interesting comparison both in camera performance but also in velocity estimation methods. The figure above shows a plot for the results of Kalman magnitude of velocity estimations for various cameras (the same cameras, used in the same order / setup for all scenes), and the corresponding ground truths from physics calculations, and radar data. This data is from an outdoor single rolling object down an inclined plane/ramp scene. As is evident from the plot, there is still much work to do in multiple aspects of velocity estimation before coming close to the either the radar or physics based ground truths.*

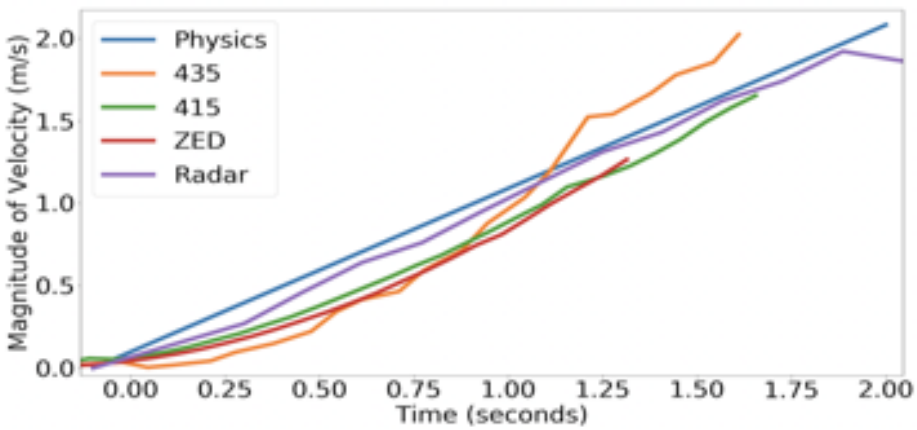

Figure 2: INDOOR RAMP SETUP - VELOCITY ESTIMATION *Above is a plot for the results of Kalman magnitude of velocity estimations for various cameras, the corresponding physics calculation, and radar data. This data is from an indoor single rolling object down an inclined plane/ramp scene from the I-MOVE dataset. of the proposed I-MOVE dataset highlights the need for improvement in the field. This plot as with 1,shows there is still much work to do before being withing the accuracy of the radar or physics based ground truths.*

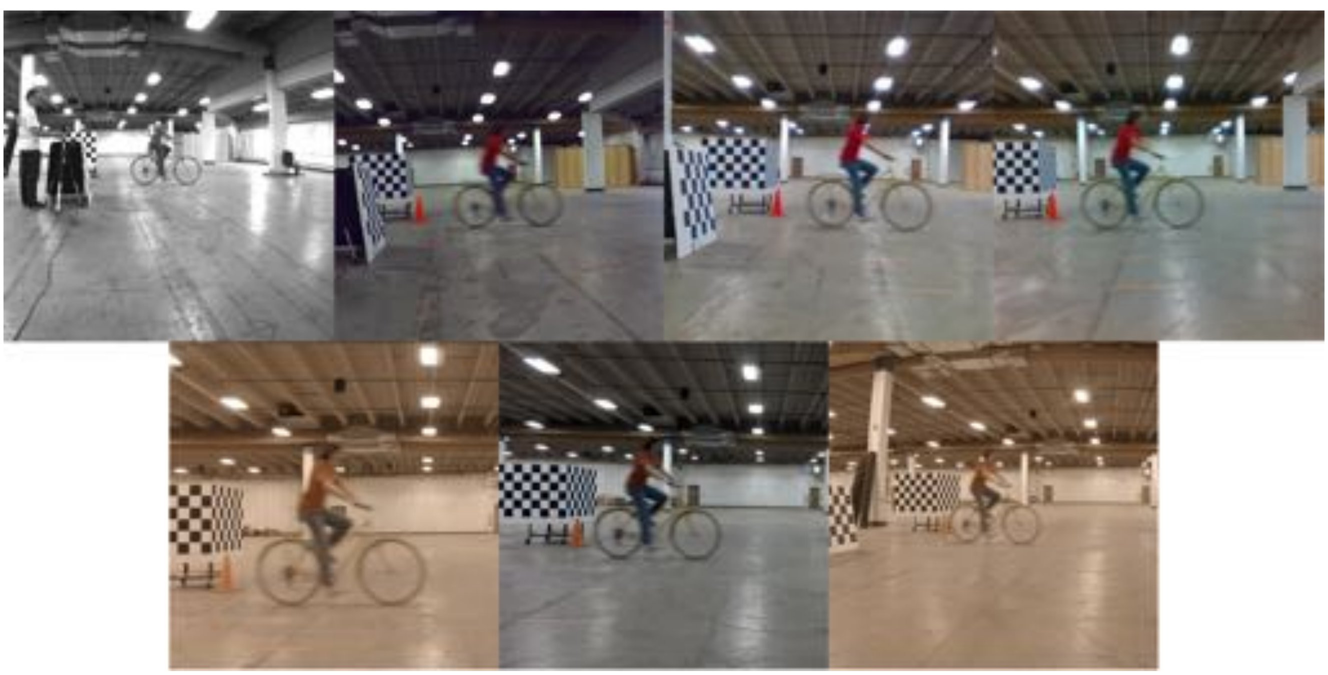

Figure 3: BIKER SAMPLE *The above figure is a frame from each device used in the I-MOVE dataset during a biking scene. The frame used from each device is the most similar timestamp / point of action. The I-MOVE dataset used seven devices for data collection to get various angles, forms of depth estimation, frame-rates, field of views, and coloration for each scene. The devices used in the top row of the figure from left to right are as follows: MYNT, ZED, Intel RealSense 415, and Intel RealSense 435. The devices used in the bottom row of the figure from left to right are all GoPros, they are as follows: modified Stereo GoPros 54, modified Stereo GoPros 64, and Standard Stereo GoPros. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

Figure 4: OBJECT DROP SAMPLE *Above is a frame from each device used in the I-MOVE dataset during a dropped object scene. The frame used from each device is the most similar timestamp / point of action, due to the fact that the cameras feature frame-rates varying from 15 - 60 frames per second there are sometimes differences in the objects state of motion and blur as can be more apparent for quicker motions or motions of smaller objects such as the above dropped object figure. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

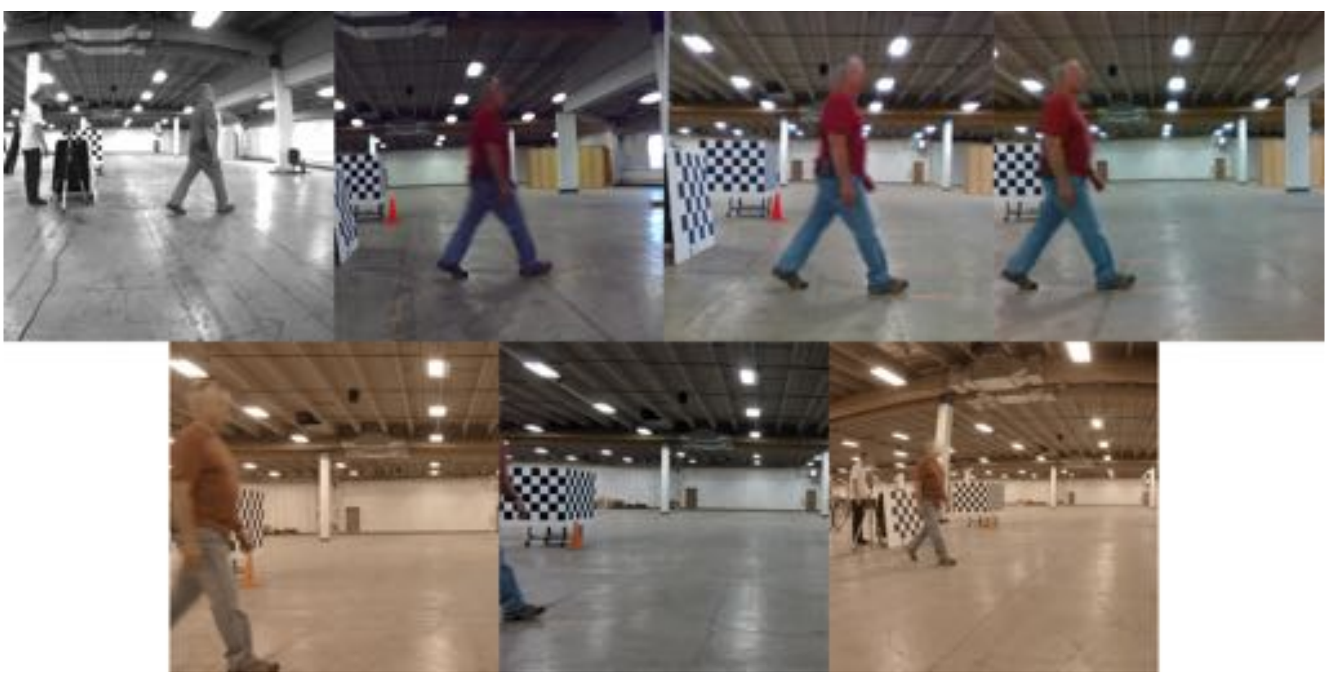

Figure 5: WALKING PEDESTRIAN SAMPLE *The figure above features a frame from each device used in the I-MOVE dataset during a walking pedestrian scene. This scene shows the large difference in angle and field of view for each device depending on where the action / object's motion is taking place. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

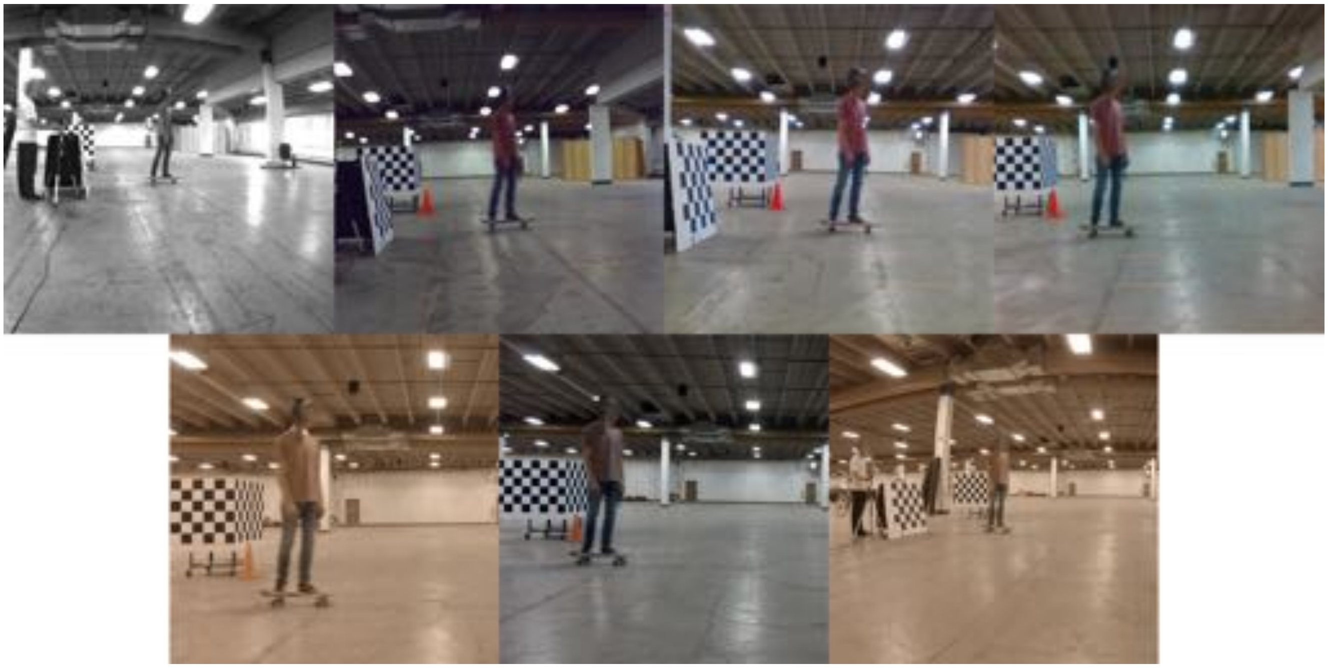

Figure 6: SKATEBOARDER SAMPLE *The I-MOVE dataset also features scenes including skateboarding, above features a frame from each device used in the I-MOVE dataset during a skateboarding scene. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

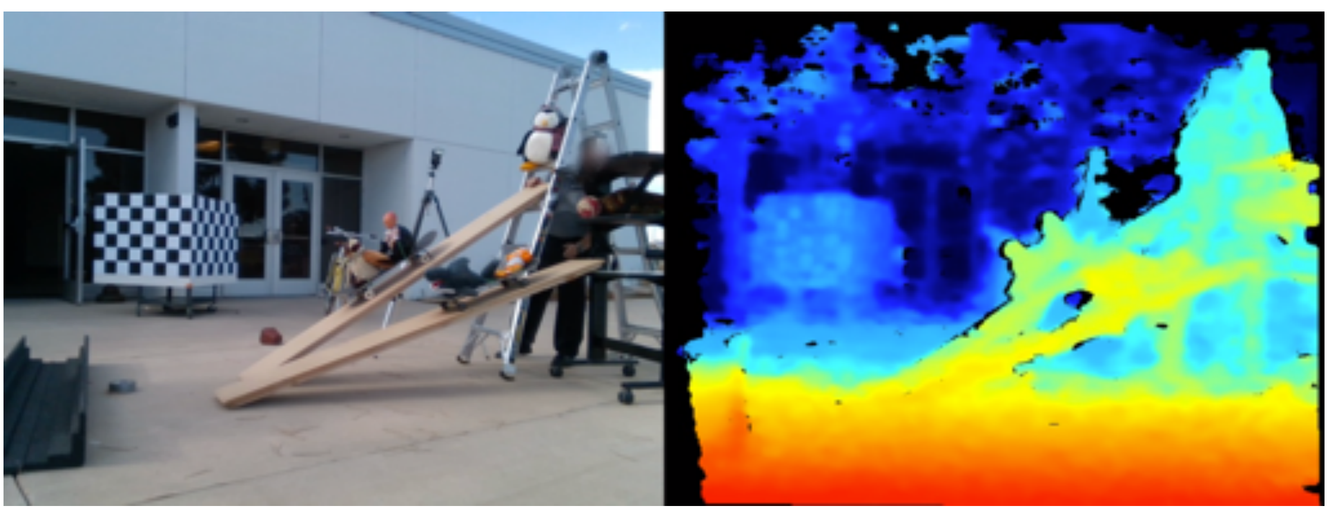

Figure 7: OUTDOOR DUAL-RAMP DEPTH SAMPLE *This figure shows a color frame and colorized depth frame from the same timestamp of a outdoor dual-ramp rolling object scene from the I-MOVE dataset. It features two objects rolling down ramps of different heights. This data was collected using the Intel RealSense 415. The accompanying video can be found in the supplemental material folder. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

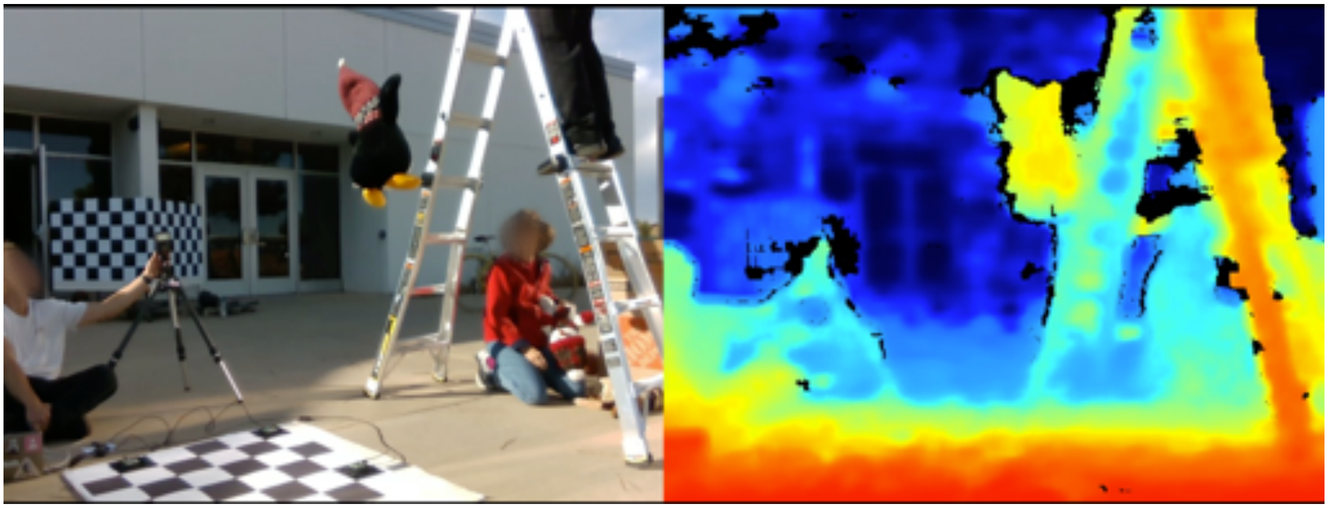

Figure 8: OUTDOOR OBJECT DEPTH DEPTH SAMPLE *The figure above is a color frame and colorized depth frame from the same timestamp of a object dropped for the I-MOVE dataset. It features multiple objects dropped from various known / documented heights. This data was collected using the Intel RealSense 415 which features a larger distance between the two cameras/lens compared to the Intel RealSense 435. The accompanying video can be found in the supplemental material folder. The faces of those involved in the making of this dataset have been blurred in an effort to follow the guidelines of the blind review.*

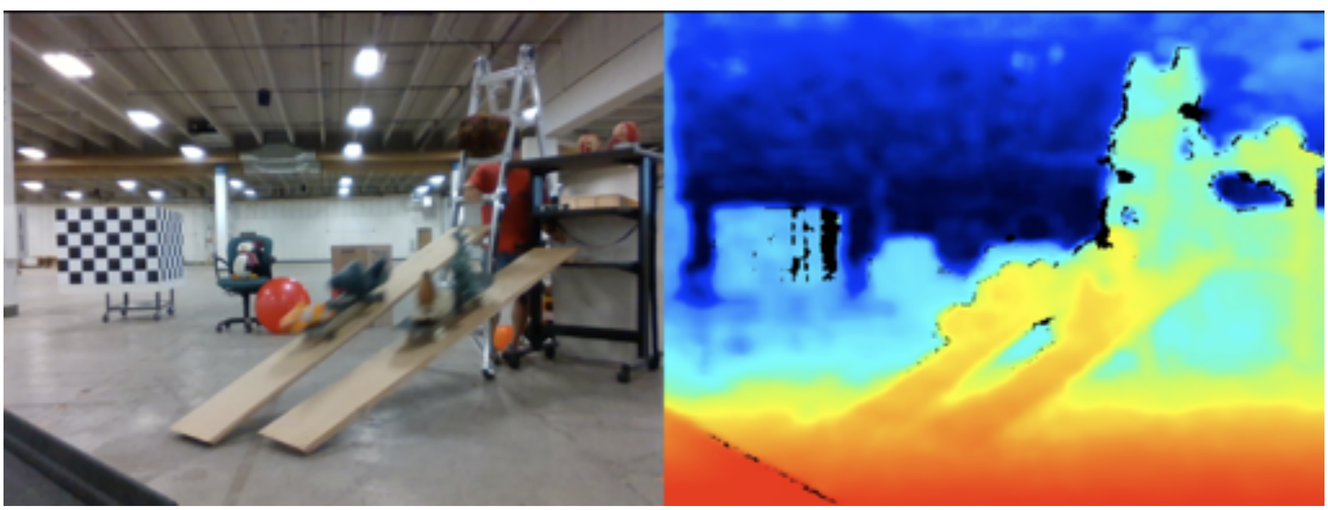

Figure 9: INDOOR DUAL-RAMP DEPTH SAMPLE *Above is a color frame and colorized depth frame from the same timestamp of a dual-ramp rolling object scene from the I-MOVE dataset. It features two objects rolling down ramps of different heights. As can be seen in this image compared to the outdoor dual-ramp data, the indoor environment causes there to be a significant more amount of motion blur. This greatly increases the difficulty of the tracking problem necessary to estimate the magnitude of velocity accurately. This data was collected using the Intel RealSense 435. The accompanying video can be found in the supplemental material folder.*