This document provides the supplementary information about the demonstration videos of the 3DROM algorithm.

1. The videos contain the pedestrian detection results of three benchmark datasets.
2. The video “3DROM_demo_Wildtrack.avi” contains the pedestrian detection results at frames 1800-1995 of the EPFL Wildtrack Dataset with 7 camera views. The Wildtrack video was originally captured at a frame rate of 60 frames per second. The demonstration video was generated by sampling the original video at a frame rate of 2 frames per second. The original resolution of the Wildtrack video is 1920 ×1080 pixels, which was reduced to 640×360 pixels in the demonstration video.
3. The video “3DROM_demo_MultiviewX.avi” contains the pedestrian detection results at frames 360-399 of the MultiviewX Dataset with 6 camera views and 40 frames in total. The original resolution of the MultiviewX video is 1920 ×1080 pixels, which was reduced to 640×360 pixels in the demonstration video.
4. The video “3DROM_demo_Terrace.avi” contains the pedestrian detection results at frames 2525-4875 of the EPFL Terrace Dataset with 4 camera views. The demonstration video was generated at the original frame rate 25 frames per second and 35 frames per second. The height and width of the frame were not changed.
5. The bounding boxes are of the average height of the pedestrians and sitting on the ground locations where the pedestrians are identified.
6. Each pedestrian is represented by a distinguished but consistent colour, at a single frame, in all camera views and the synthetic top view. This colour for each pedestrian may change across different frames, since tracking is not included in the 3DROM algorithm.
7. The red rectangle, on the ground plane in each video dataset, represents the Area of Interest (AOI) region. Only the pedestrians within the AOI are detected in each dataset. Some pedestrians, who are partially occluded by the lower border of a camera view, are missed in the detection, since their feet are out of the AOI.