This document provides the supplementary information about the demonstration videos of the TMVD method.

1. The video "Demo_WILDTRACK.avi" contains the pedestrian detection results at frames 1800-1995 of the 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.

2. The video "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.

3. The bounding box was generated by back-projecting the eight vertices of a cuboid from 3D space to each camera view and finding the minimum rectangular box which contains all the eight projected vertices. This cuboid sits at the location of that detected pedestrian on the ground plane and is of the average pedestrians' size.

4. Each pedestrian is represented by a distinguished but consistent colour, at a single frame, in all camera views and a virtual top view. The colour for the same pedestrian may change across different frames, since tracking is not included in the proposed method.

5. 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.