This document provides the supplementary information about the QM demonstration video.

1. The algorithm used to produce the video is QM + DeepLab.

2. The video contains the pedestrian detection results in part of the frames of the EPFL Terrace Dataset with 4 camera views, which is to make the size of the video smaller.

3. The bounding boxes are of the average size of the pedestrians and sitting on the ground locations where the pedestrians are identified. Identified phantoms are not displayed to avoid confusion.

4. Each pedestrian is represented by a distinguished but consistent colour, at a single frame, in the four camera views and the synthetic top view. This colour may change across different frames.