The workflow is demonstrated in Fig. 1. This is an unsupervised method based on supervoxels and the multi-level semantic relationships. The point clouds are first filtered into ground and non-ground points using our adaptive surface filter; then the non-ground points are classified. We model the classification problem as a high order Markov Random Field (MRF) as below,

$$E(\boldsymbol{y}) = \sum_{i \in \boldsymbol{v}} E(y_i) + \sum_{(i,j) \in \boldsymbol{e}} E(y_i, y_j) + \sum_{c \in \boldsymbol{c}} E(y_c) \quad (1)$$

where $\boldsymbol{v}$, $\boldsymbol{e}$ and $\boldsymbol{c}$ are first order, second order and high order cliques in the graphical model, respectively. Because minimizing the energy of high order MRF is hard, we separate it into two steps: a) using graph cut with first and second order cliques to get an initial classification and then b) refine them using the high order energy in a greedy way. At last, the classified non-ground points are merged with the ground points for the final results. Buildings are extracted from the classified point clouds.

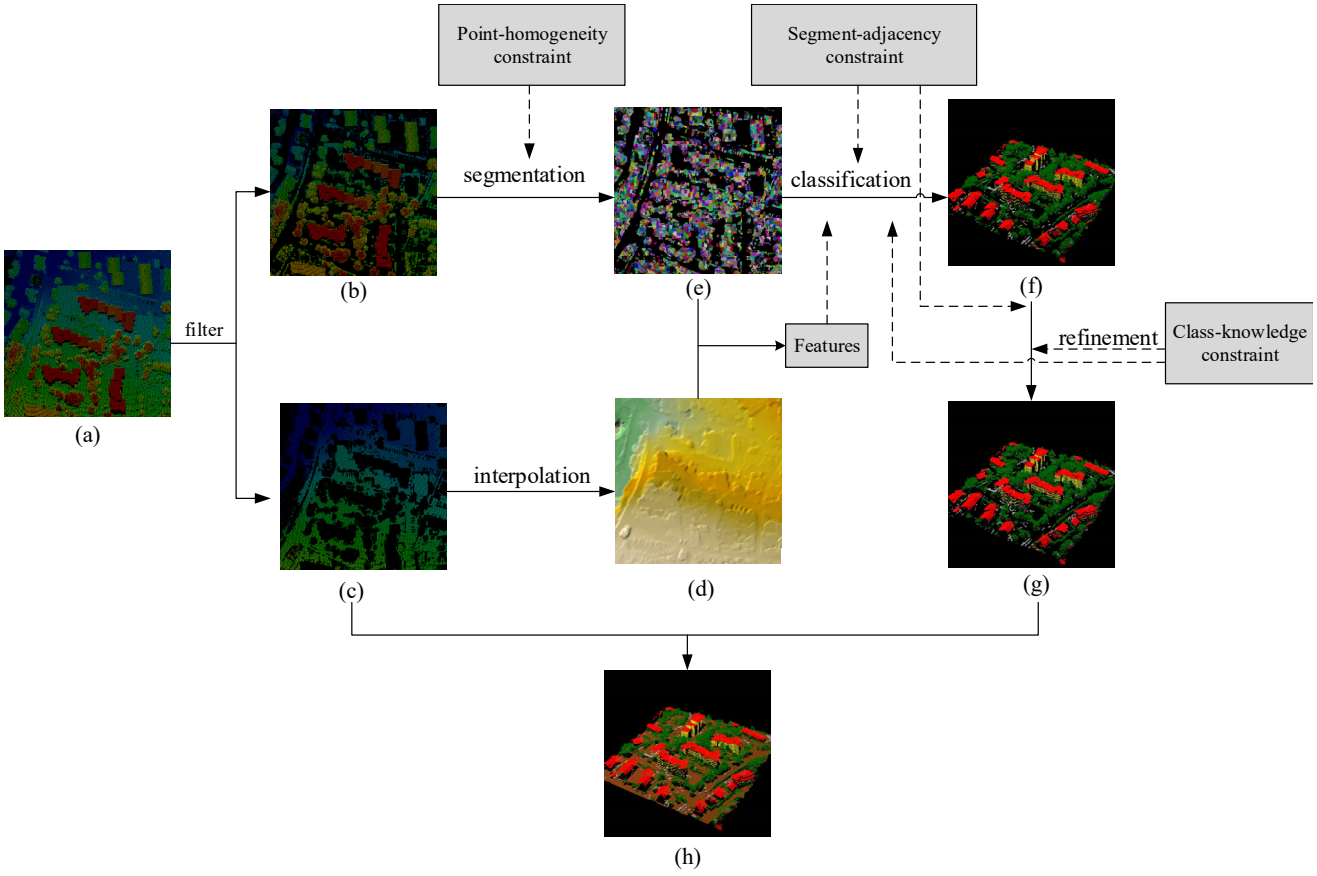

Fig. 1 Workflow of the framework. (a) Original point cloud, (b) Non-ground points set, (c) Ground points set, (d) DEM interpolated from (c), (e) supervoxels, (f) Initial classes, (g) Refined classes, (g) Final classes: roofs, facades, vegetation, clutters and ground.