Brief description of the results

1) The results are presented as geo-referenced white-black images, in which white pixels mean buildings while as back pixels mean backgrounds.
2) The three images building_geo_1,2,3 are results for Vaihingen test area 1,2,3 respectively.
3) They are based on three small images covering area1,2,3, and they are cut from the TOP_Mosaic_09cm.tif image. In other words, we shift the up-left corner and shorten the width and height (pixels) to get the small images. The coordinates of the up-left corner and the size (0.09 m) of each pixel are recorded in the geotiff file for the results.

Brief description of the method

Building extraction is formulated as a foreground and background segmentation problem, which in turn is formulated as a pixel labeling problem. Each pixel has one label denoting its class, i.e. foreground or background. An optimal label field is estimated from observed data based on a Conditional Random Field describing the posterior of label field given observed data.

There are some existing methods on building extraction based on Conditional Random Field. Some deal with building extraction from images while as others deal with building extraction from point cloud. In this paper, we put aerial images and point cloud (in the form of DSM) together and present a framework for building extraction from multiple cues, i.e. color, height and edge.

Color from images and height from DSMs are used to construct associate potential. To measure color distribution, two Gaussian Mixture Models (GMMs) are constructed for the foreground and the background areas respectively, which are learned from real data. A uniform distribution is adopted to measure the height distribution of objects. Ground areas are distinguished from off-ground areas using a flood filling schema, and then approximate heights of objects are calculated by subtracting the ground heights, and then their height are used to calculated the height probabilities. The probabilities correspond to colors and heights are fused to define the associate potentials of the CRF.

Edges from images and DSMs are used to construct the interaction potential. First, the edges are detected from images and DSMs respectively. Then, they are fused to define the potentials between neighboring pixels. Pixels across edges have lower potential than pixels not across edges.

Graph cut algorithm is applied to find the optimal configuration corresponding to the minimal energy.