

Change Detection in Urban Streets by a Real Time Lidar Scanner and MLS Reference Data

Bence Gálai¹ and Csaba Benedek^{1,2}

¹Machine Perception Research Laboratory, Institute for Computer Science and Control
H-1111 Kende u. 13-17 Budapest, Hungary

²Péter Pázmány Catholic University, H-1083, Práter u. 50/A, Budapest, Hungary
{galai.bence, benedek.csaba}@sztaki.mta.hu

Abstract. In this paper, we introduce a new technique for change detection in urban environment based on the comparison of 3D point clouds with significantly different density characteristics. Our proposed approach extracts moving objects and environmental changes from sparse and inhomogeneous instant 3D (i3D) measurements, using as reference background model dense and regular point clouds captured by mobile laser scanning (MLS) systems. The introduced workflow consist of consecutive steps of point cloud classification, crossmodal measurement registration, Markov Random Field based change extraction in the range image domain and label back projection to 3D. Experimental evaluation is conducted in four different urban scenes, and the advantage of the proposed change detection step is demonstrated against a reference voxel based approach.

Keywords: Change detection, Lidar

1 Introduction

The progress of real time Lidar sensors, such as rotating multi-beam (RMB) Lidar scanners, open several new possibilities in comprehensive environment perception for autonomous vehicles (AV) and mobile city surveillance platforms. On one hand, RMB Lidars directly provide instant 3D (i3D) information facilitating the detection of moving street objects and environmental changes. On the other hand, with registering the i3D measurements to a detailed 3D city map, the detected objects and changes can be accurately localized and mapped to a geo-referred global coordinate system.

Using new generation Geo-Information Systems, several major cities maintain from their entire road network dense and accurate 3D point cloud models obtained by Mobile Laser Scanning (MLS) technology. As a possible future utilization, these MLS point clouds can be efficiently considered by the AV's onboard i3D environment sensing modules as highly detailed reference background models. In this context, *change detection* between the instantly sensed RMB Lidar measurements and the MLS based reference environment model appears as a crucial task, which indicates a number of key challenges.

Particularly, there is a significant difference in the quality and the density characteristics of the i3D and MLS point clouds, due to a trade-off between temporal and spatial

resolution of the available 3D sensors. RMB Lidar scanners, such as the Velodyne HDL-64 provide sequences of full-view point cloud frames with 10-15fps, and the size of the transferable data is also limited enabling real time processing. As a consequence the measurements have a low spatial density, which quickly decreases as a function of the distance from the sensor, and the point clouds may exhibit particular patterns typical to sensor characteristic, such as the ring patterns of the Velodyne sensor (see Fig. 1(c)). Although the 3D measurements are quite accurate (up to few cms) in the sensor's local coordinate system, the global positioning error of the vehicles may reach several meters in city regions with poor GPS signal coverage.

Recent MLS system such as the Riegl VMX450 are able to provide dense and accurate point clouds from the environment with homogeneous scanning of the surfaces (Fig. 1(a)-1(b)) and a nearly linear increase of points as a function of the distance. The point density of MLS point clouds is with 2-3 orders of magnitude higher than the density of i3D scans which makes direct point-by-point comparison inefficient. On the other hand, due to the sequential environment scanning process, the result of MLS is a static environment model, which can be updated typically with a period of 1-2 years in large cities. Therefore, apart from the changes caused by moving objects we must expect various differences caused by environmental changes such as altering the buildings and street furniture, or seasonal changes of the tree-crowns or bushes etc.

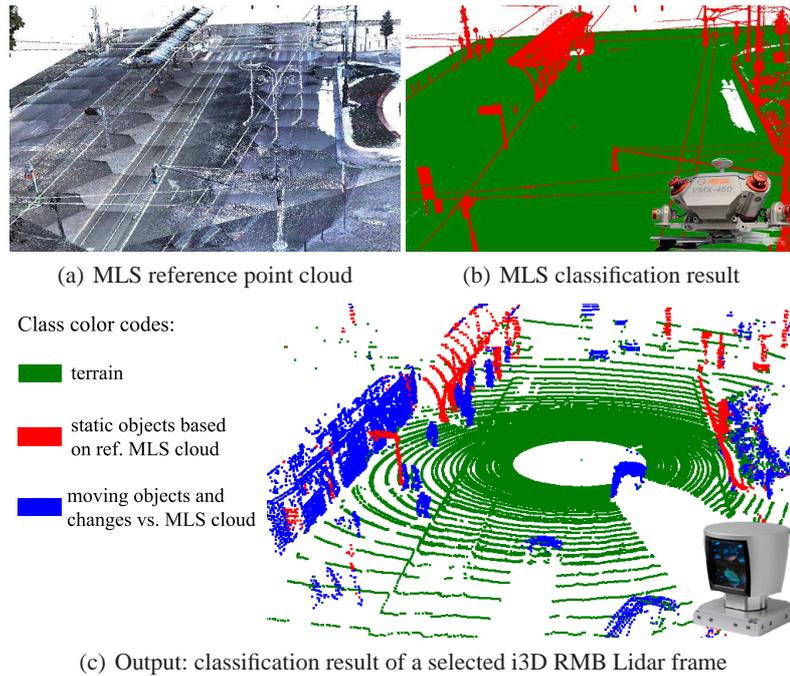


Fig. 1. Overview on the proposed approach: based on reference MLS data (a,b), the goal is separation of static scene elements and moving objects/changes on instant RMB Lidar frames (c)

2 Previous Work

In the recent years various techniques have been published for change detection in point clouds, however, the majority of the approaches rely on dense terrestrial laser scanning (TLS) data recorded from static tripod platforms [1, 2]. As explained in [1], classification based on calculation of point-to-point distances may be useful for homogeneous TLS and MLS data, where changes can be detected directly in 3D. However, the point-to-point distance is very sensitive to varying point density, causing degradation in our addressed i3D/MLS cross-platform scenario. Instead, [1] follows a ray tracing and occupancy map based approach with estimated normals for efficient occlusion detection, and point-to-triangle distances for more robust calculation of the changes. Here the Delaunay triangulation step may mean a critical point, especially in noisy and cluttered segments of the MLS point cloud, which are unavoidably present in a city-scale project. [2] uses a nearest neighbor search across segments of scans: for every point of a segment they perform a fixed radius search of 15 cm in the reference cloud. If for a certain percentage of segment points no neighboring points could be found for at least one segment-to-cloud comparison, the object is labeled there as moving entity. A method for change detection between MLS point clouds and 2D terrestrial images is discussed in [3]. An approach dealing with purely RMB Lidar measurements is presented in [4], which use a ray tracing approach with nearest neighbor search. A voxel based occupancy technique is applied in [5], where the authors focus on detecting changes in point clouds captured with different MLS systems. However, the differences in data quality of the inputs are less significant than in our case.

3 Proposed Change Detection Method

We assume that the reference MLS data is accurately geo-referred, and the i3D Lidar platform also has a coarse estimation of its position up to maximum 10 m translational error. Initially, the orientation difference between the car’s local and the MLS point cloud’s global coordinate systems may be arbitrarily large (see Fig. 2). The proposed approach consists of four main steps: ground removal by point cloud classification, i3D–MLS point cloud registration, change detection in the 2D range image domain, and label backprojection to the 3D point cloud.

The *ground removal* step separates terrain and obstacle regions using a locally adaptive terrain modeling approach, expecting inhomogeneous RMB Lidar point clouds with typically non-planar ground. First we fit a regular 2D grid with fixed rectangle side length onto the horizontal $P_{z=0}$ plane, using the Lidar sensor’s vertical axis as the z direction. We assign each p point of the point cloud to the corresponding cell, which contains the projection of p to $P_{z=0}$. After excluding the sparse grid cells, we use point height information for assigning each cell to the corresponding cell class. All the points in a cell are classified as ground, if the difference of the minimal and maximal point elevations in the cell is smaller than an elevation threshold (used 25cm), moreover the average of the elevations in neighboring cells does not exceeds an allowed height range. The result of ground segmentation is shown in Fig. 1(b) and 1(c), which confirms that our technique handles robustly the various i3D and MLS Lidar point cloud types.

For *point cloud registration* we adopt our latest technique [6] for matching point cloud measurements with significantly different density characteristics. The registration process includes three steps. *First*, following the removal of ground points, we search for distinct groups of close points in the remaining obstacles cloud, and assign each group to an abstract object. For handling difficult scenarios with several nearby adjacent objects, we adopted a hierarchical 2-level model [7], which separates first large objects or object groups at a coarse grid level with large cells, then in the refinement it can efficiently separate the individual objects within each group. *Second*, we coarsely align the two point clouds by considering only the center points of the previously extracted abstract objects. We apply here the generalized Hough transform to extract the best similarity transformation in the sense that when applying the transformation to the object centers in the first frame as many of these points as possible overlap with the object centers in the second frame [6]. *Third*, we run a point-level refinement on the above approximate global transform, applying the Normal Distribution Transform (NDT) for all object points. The success of the registration process from an extremely weak initial point cloud alignment is demonstrated in Fig. 2.

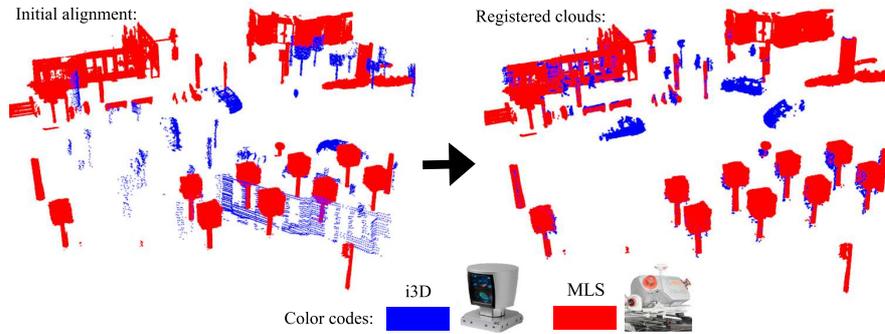


Fig. 2. Demonstration of the proposed point cloud registration step (Deák tér, Budapest). Blue and red points represent the i3D and MLS point clouds, respectively.

The *change detection* module receives a co-registered pair of i3D and MLS point clouds, where the terrain is already removed (see Fig. 2 right image). Our proposed solution extracts changes in the range image domain. Creating a range image I_{i3D} from the RMB Lidar’s point stream is straightforward as its laser emitter and receiver sensors are vertically aligned, thus every measured point has a predefined vertical position in the image, while consecutive firings of the laser beams define their horizontal position. Geometrically, this mapping is equivalent to projecting the 360° obstacle point cloud to a cylinder surface, whose main axis is equal to the vertical axis of the RMB Lidar scanner. Using Velodyne HDL-64 sensor with 15 Hz rotation frequency, the typical size of this I_{i3D} range image is 64×1024 . Since the above projection only concerns the obstacle cloud (without the ground), and several fired laser beams do not produce reflections at all (such as those from the direction of the sky), several pixels of

the range map will be assigned to zero (i.e. invalid) depth values. Moreover, such holes may also appear in the range maps due to noise or quantization errors of the rotation angles. On account of this artifact we interpolate the pixel values which have in their 8-neighborhood at least four valid (non-zero) neighboring depth values, as demonstrated in Fig. 3. A sample full-view i3D range image is shown in Fig. 4(a).



Fig. 3. Range image segment from the Velodyne i3D sensor

The reference background range image is generated from the 3D MLS point cloud with ray tracing, exploiting that the current position and orientation of the RMB Lidar platform are available in the reference coordinate system as a result of the point cloud registration step. Thereafter simulated rays are emitted into the MLS cloud from the moving platform's center position with the same vertical and horizontal resolution as the RMB Lidar scanner. To handle minor registration issues and sensor noise, each range image pixel value is determined by examining multiple MLS points lying inside a pyramid around the simulated RMB Lidar ray. For a given pixel of the MLS range map the depth values of the corresponding points are weighted with a sigmoid function:

$$I_{\text{MLS}}(i, j) = \frac{\sum_{k=1}^{K^{i,j}} w_k^{i,j} D_k^{i,j}}{\sum_{k=1}^{K^{i,j}} w_k^{i,j}}, \quad w_k^{i,j} = \frac{1}{1 + e^{l(D_k^{i,j} - \min_{i,j} D_k^{i,j}) - m}}, \quad (1)$$

where $K^{i,j}$ is the number of MLS points in the (i, j) pyramid, $D_k^{i,j}$ is distance of the k -th point from the ray origin, and the weights $w_k^{i,j}$ are calculated using a sigmoid function ($l = 0.5$ and $m = 5$ parameters were empirically set). This calculation formula ensures that the nearest points within the pyramid receive the highest weights, but due to the smoothing effect of weighted averaging, the presence of outlier points, or highly scattered regions (such as vegetation) do not cause significant artifacts. A sample MLS range image generated by the above process is shown in Fig. 4(b).

In the next step, the calculated RMB Lidar-based I_{i3D} , and MLS-based I_{MLS} range images are compared using a Markov Random Field (MRF) model, which classifies each pixel of the range image lattice as foreground (FG) or background (BG). Foreground pixels represent either moving/mobile objects in the RMB Lidar scan, or various environmental changes appeared since the capturing date of the MLS point cloud.

Two sigmoid functions are used to define fitness scores for each class:

$$F_{\text{BG}}(i, j) = 1 - \frac{1}{1 + e^{(d^{i,j} - a^{i,j})}}, \quad F_{\text{FG}}(i, j) = 1 - \frac{1}{1 + e^{-(d^{i,j} - a^{i,j})}}, \quad (2)$$



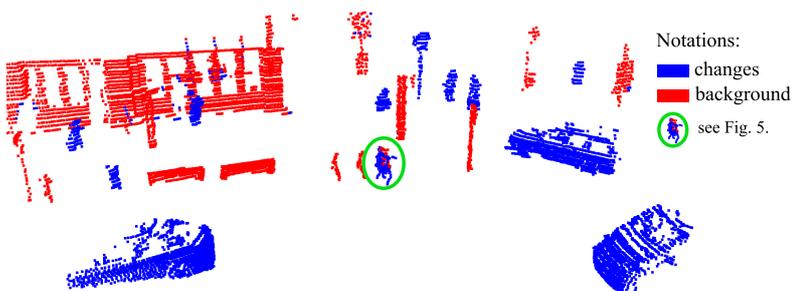
(a) Filtered & interpolated Velodyne (i3D) range image



(b) MLS based range image from the actual position of the Velodyne platform

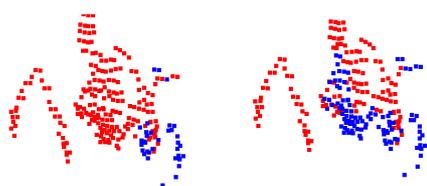


(c) MRF based change mask in the range image domain



(d) Back projection of the change mask to the Velodyne point cloud

Fig. 4. Demonstration of the proposed MRF based change detection process in the range image domain, and result of label back projection to the 3D point cloud



(a) Voxel based method (b) Proposed method

Fig. 5. Comparison of the voxel based reference and the proposed range image based approach: a sample bike shed from a magnified image part of the scene in Fig. 4.

where $d^{i,j} = I_{i3D}(i, j)$ and $a^{i,j} = I_{MLS}(i, j)$.

To formally define the range image segmentation task, we assign to each (i, j) pixel of the pixel lattice S a $l_{i,j} \in \{FG, BG\}$ class label so that we aim to minimize the following energy function:

$$E = \sum_{(i,j) \in S} V_D(d^{i,j}|l_{i,j}) + \sum_{(i,j) \in S} \sum_{(m,n) \in N_{i,j}} \beta \cdot 1\{l_{i,j} \neq l_{m,n}\}, \quad (3)$$

where $\beta > 0$ is a smoothness parameter for the label map (used $\beta = 0.5$), and $N_{i,j}$ the four-neighborhood of pixel (i, j) . $V_D(d^{i,j}|l_{i,j})$ denotes the data term, derived as:

$$V_D(d^{i,j}|l_{i,j} = BG) = -\log(F_{BG}(i, j)), \quad V_D(d^{i,j}|l_{i,j} = FG) = -\log(F_{FG}(i, j))$$

The MRF energy (3) is minimized via the fast graph-cut based optimization algorithm [8], which process results in a binary change mask in the range image domain, as shown in Fig. 4(c). The final step is *label backprojection* from the range image to the 3D point cloud (see Fig. 4(d)), which can be performed in a straightforward manner, since in our i3D range image formation process, each pixel represents only one Velodyne point.

4 Experiments

We have evaluated the proposed change detection technique in four test scenarios. Each test sequence contains 70 consecutive time-frames from the RMB Lidar sensor, where each i3D frame has a GPS-based coarse location estimation for the point cloud centers, with maximum few meters position error. The MLS reference cloud is accurately georeferenced, and we assume that it only contains the static scene elements such as roads, building facades, and street furniture. For each RMB Lidar frame, we execute the complete workflow of the proposed algorithm.

The Ground Truth (GT) labeling of the RMB Lidar’s i3D point clouds was done in a semi-automatic manner. First, using the registered i3D and MLS frames, we applied an automated nearest neighbor classification with a small distance threshold (3 cm), thereafter the labeling of the changed regions was manually revised. As evaluation metrics, we calculated the Precision, Recall and F score values of the detection output at point level, based on comparison to the GT.

Since we have not found any similar i3D-MLS crossmodal change detection approach in the literature, we adopt a voxel based technique [5] as reference, which was originally constructed for already registered MLS/TLS point clouds. Therefore by testing both the proposed and the reference models, we apply the same registration workflow introduced in Sec. 3, and only compare the performance of the voxel based and the proposed range image based change detection steps. The reference voxel based technique fits a regular 3D voxel grid to the registered point clouds, thereafter a given RMB Lidar point is classified as foreground if and only if its corresponding voxel does not contain any points in the MLS cloud. We tested this method with multiple w voxel sizes, which parameter naturally affects both the detection performance and the computational time. With larger voxels, we cannot detect some changes in cluttered regions, where the objects can be close to each other and to various street furniture elements. On the other

Table 1. Quantitative comparison of the Voxel based (VOX) and the proposed MRF-range image based (MRF) methods on the four test scenes, considering all regions (left), and only the crowded sidewalk areas (right)

Scenes		Overall test set			Sidewalk areas only		
		Precision	Recall	F score	Precision	Recall	F score
Deák	VOX	0.99	0.87	0.93	0.81	0.71	0.76
	MRF	0.99	0.90	0.94	0.87	0.89	0.88
Astoria	VOX	1.00	0.94	0.97	0.88	0.81	0.84
	MRF	0.95	0.98	0.97	0.84	1.00	0.91
Kálvin	VOX	1.00	0.94	0.97	0.89	0.96	0.92
	MRF	1.00	0.97	0.98	0.87	0.99	0.93
Fővám	VOX	0.98	0.70	0.82	0.84	0.64	0.73
	MRF	0.94	0.83	0.88	0.81	0.97	0.88

hand, maintaining and processing a fine 3D grid structure with small voxels requires more memory and processing time. The results shown in the upcoming comparative experiments correspond to the voxel size $w = 30$ cm, since we observed with this parametrization approximately the same running speed as using our proposed MRF-range image based model: the change detection step in each frame takes here around 80 msec on a desktop computer, with CPU implementation. Note that by decreasing the w parameter to 20 cm and 10 cm, respectively, the calculation time of the voxel based model starts to rapidly increase (120 msec and 510 msec/frame, resp.), without significant performance improvements.

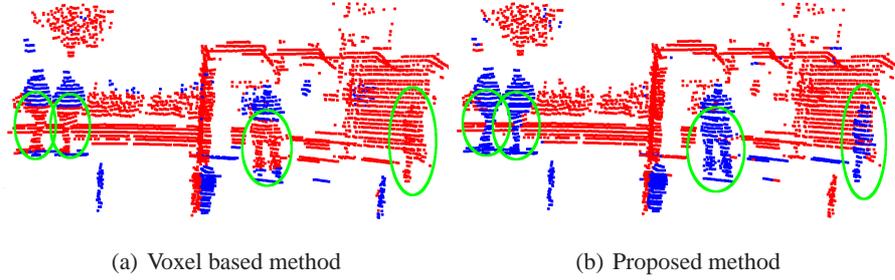


Fig. 6. Results for a sample region captured at Fővám tér, Budapest, by (a) the voxel based approach, (b) the proposed method. Red and blue points represent the detected background and foreground points respectively. Differences are marked with green ellipses.

The comparative results considering the complete dataset are shown in Table 1 (left section), which confirms that the proposed method has an efficient overall performance, and it outperforms the voxel based method in general with 1-6% F scores in the different scenes. We have experienced that the main advantage of the proposed technique

is the high accuracy of change detection in cluttered street regions, such as sidewalks with several nearby moving and static objects. As shown in Table 1 (right section), if we restrict the quantitative tests to the sidewalk areas, our method surpasses the voxel approach with 7-15% gaps in three scenes. Similar trends can be observed from the qualitative results of Fig. 5. and 6, which show successful detection samples of small object segments and fine changes with our proposed method, and corresponding limitations of the voxel based approach. As shown in Fig. 6. the voxel based technique results in many falsely ignored moving object segments, in particularly in the regions where people were standing next to static objects. On the other hand, vehicles on the roads with relatively large distances from the street furniture elements can be well separated even with large voxels, therefore the difference between the two methods is less significant in the road regions of the test scenes. Fig. 7 shows another test scene.

We display in Fig. 8 synthesized view, visualizing the point clouds of moving objects detected by the i3D RMB Lidar over the geo-referred MLS background data¹.

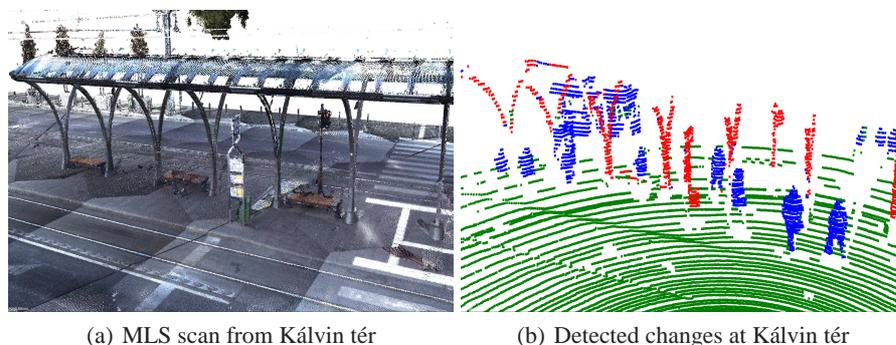


Fig. 7. *Left:* MLS laser scan of a tram stop in Kálvin tér, Budapest. *Right:* Detected changes at the tram stop. Red, blue and green points represent background objects, foreground objects and ground regions, respectively.

5 Conclusion and Future Work

We introduced a new method for change detection between different laser scanning measurements captured at street level. The results show that even small and detailed changes can be observed with the proposed method, which cannot be achieved with voxel based techniques. Future work will present a deeper investigation of various background change classes, and tests with lower resolution Lidar sensors.

Acknowledgment This work was supported by the Hungarian National Research, Development and Innovation Fund (NKFIA #K-120233). C. Benedek also acknowledges

¹ Demo video: http://web.eee.sztaki.hu/i4d/demo_iciar17.html

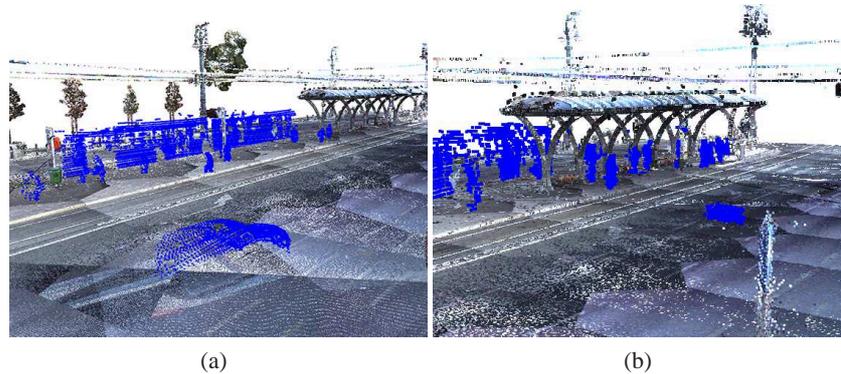


Fig. 8. Synthesized view for demonstrating geo-referred moving object detection: object point clouds (tram, car, pedestrians) detected on two subsequent i3D Velodyne frames (marked with blue) are put in and displayed in the MLS reference point cloud

the support of the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. MLS test data was provided by the Road Management Department of the City Council of Budapest (Budapest Közút Zrt).

References

1. W. Xiao, B. Vallet, M. Brédif, and N. Paparoditis, “Street environment change detection from mobile laser scanning point clouds,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 107, pp. 38–49, 2015.
2. A. Schlichting and C. Brenner, “Vehicle localization by lidar point correlation improved by change detection,” *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B1, pp. 703–710, 2016.
3. R. Qin and A. Gruen, “3D change detection at street level using mobile laser scanning point clouds and terrestrial images,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 90, pp. 23–35, 2014.
4. J. P. Underwood, D. Gillsjö, T. Bailey, and V. Vlaskine, “Explicit 3D change detection using ray-tracing in spherical coordinates,” in *IEEE International Conference on Robotics and Automation*, (Karlsruhe, Germany), pp. 4735–4741, May 2013.
5. K. Liu, J. Boehm, and C. Alis, “Change detection of mobile LIDAR data using cloud computing,” *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLI-B3, pp. 309–313, Jun 2016.
6. B. Gálai, B. Nagy, and C. Benedek, “Crossmodal point cloud registration in the Hough space for mobile laser scanning data,” in *IEEE International Conference on Pattern Recognition (ICPR)*, (Cancun, Mexico), 2016.
7. A. Börcs, B. Nagy, and C. Benedek, “Fast 3-D urban object detection on streaming point clouds,” in *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving at ECCV’14*, vol. 8926 of LNCS, pp. 628–639, Zürich, Switzerland: Springer, 2015.
8. Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.