LEGION SEGMENTATION FOR BUILDING EXTRACTION FROM LIDAR BASED DSM DATA

Chun Liu[a,b], Beiqi Shi[a,c,*], Xuan Yang[a] and Nan Li[a]

[a]Department of Survey and Geo-Informatics, Tongji University, Shanghai, China, 200092 liuchun@tongji.edu.cn

[b]Key Laboratory of Advanced Engineering Surveying of NASMG, Shanghai, China, 200092

[c]Urban Information Research Center, Shanghai Normal University, Shanghai, China, 200234 carashi@163.com

Commission III, WG III/4

KEY WORDS: LiDAR DSM, LEGION segmentation, Building Extraction, Height Texture

ABSTRACT:

Recently, a neural oscillator network based on biologically framework named LEGION (Locally Excitatory Globally Inhibitory Oscillator Network),which each oscillator has excitatory lateral connections to the oscillators in its local neighbourhood as well as a connection with a global inhibitor, has been applied to segmentation field. The extended LEGION approach is constructed to extract buildings digital surface model (DSM) generated from LiDAR data. This approach is with no assumption about the underlying structures in DSM data and no prior knowledge regarding the number of regions. Instead of using lateral potential to find a major oscillator block in original way, Gray Level Co-occurrence Matrix (GLCM) homogeneity measuring DSM height texture is applied to distinguish buildings from trees and assist to find LEGION leaders in building targets. Alongside the DSM height texture attribute, extended LEGION can extract buildings close to trees automatically. Then a solution of least squares with perpendicularity constraints is put forward to determine regularized rectilinear building boundaries, after tracing and connecting the rough building boundaries. In general, the paper presents the concept, algorithms and procedures of the approach. It also gives experimental result of Vaihingen A2 region by then ISPRS test project and another result based on a DSM data of suburban area. The experiment result showed that the proposed method can effectively produce more accurate buildings boundary extraction.

1. Introduction

Building representations are needed in a variety of applications, such as cartographic analysis, urban planning, and visualization. And the development of building automated extraction algorithms is of great importance. Since LiDAR is a fast method for sampling the earth's surface with a high density and high point accuracy, many attempts have been made on building extraction from a digital surface model (DSM) generated from LiDAR data. Wang and Schenk (2000) generate the triangulated irregular network (TIN) model from the LiDAR point clouds. Triangles are then grouped based on the orientation and position to form larger planar segments. The intersection of such planar segments results in building corners or edges. Al-Harthy and Bethel (2002) determine the building footprints by subtracting DTM from DSM obtained by initially filtering out the non-ground points. The building polygon outline is then obtained by using a rotating template to determine the angle of highest cross-correlation, which suggests the dominant directions of the building. Miliaresis and Kokkas (2007) presented a new method for the extraction of a class for buildings from LiDAR DEMs on the basis of geomorphometric segmentation principles. It is difficult to remove vegetation in urban or suburban areas. Most popular approaches were to detect buildings by fusing LiDAR data with multi-spectral images(Walter ,2004; Lu et al., 2006; Li et al., 2010). However, fusing LiDAR data with multi-spectral data with different resolutions may add errors to building detection (Tullis and Jensen, 2003), the purpose of this paper is to develop an alternative automatic building extraction method based only on LiDAR data.

We use an extended neural oscillator network approach for segmenting LiDAR DSM imagery into semantically meaningful entities and extracting buildings objects. This is based on temporal correlation theory to address the binding problem by using a biologically plausible representation. The process consists of a sequence: After generating DSM, a neural oscillator network based on biologically framework named LEGION segmentation is constructed and applied to extract buildings from DSM; the rough building boundaries are traced and connected; in the final step, all boundary points are integrated in a least squares solution with perpendicularity constraints to determine a regularized rectilinear boundary. The experiment on the given data provided by the "ISPRS Test Project on Urban Classification and 3D Building Reconstruction" verified that the proposed method can produce accurate buildings boundary extraction.

2. Methodology

The purpose of LiDAR DSM segmentation is to separate the DSM data into different classes depending on specific application requirements, such as building extraction. The region segmentation method, using height differences between neighboring grid points checked against a predetermined threshold value, has difficulty in segmenting DSM into semantically meaningful entities. The neural oscillator network

* Corresponding author: Beiqi Shi, Ph.D. candidate, research field in neural oscillator network and its application

approach has several major advantages over other methods on DSM data. Firstly, segmentation is achieved mainly by local computation. Secondly, unlike other artificial neural network approaches, it does not ask for training data. Thirdly, this neural oscillator network need not always produce periodic behavior (Wang, 2007). Finally, the neural oscillator network approach is a dynamic system with parallel and neutrally implementable computation (Wang and Terman, 1997). Thus, we use LEGION scheme to address the problem with no assumption about the underlying structures in DSM data and no prior knowledge regarding the number of regions to extract building objects.

LEGION is a network of Terman-Wang oscillators which comprise a large class of nonlinear dynamic systems, and arise naturally from neuron-physiological systems(Wang and Terman, 1995). Based on temporal correlation theory, LEGION can address the binding problem by using a biologically plausible representation. Each oscillator in the LEGION network connects excitatorily with the oscillators in its neighborhoods as well as inhibitorily with a global inhibitor.

2.1 Original LEGION Algorithm

The basic unit of LEGION is a relaxation oscillator defined as a feedback loop between an excitatory variable xi and an inhibitory yi, where x-nullcline is a cubic function and the y-nullcline is a sigmoid function. It is described as follow.

$$x_i' = 3x_i - x_i^3 + 2 - y_i + I_i H(p_i - \theta) + S_i + \rho \qquad (1)$$
$$y_i' = \varepsilon(\gamma(1 + \tanh(x_i / \beta)) - y_i)$$

In this formula, Ii represents external stimulation to the oscillator. $H(p_i-\theta)$, a Heaviside function, distinguishes a major oscillator block to address the fragmentation problem. p_i is the potential of the oscillator i and θ is a threshold, where $0<\theta<1$. ρ denotes the amplitude of Gaussian noise. ε defines a typical relaxation oscillator with two time scales. The parameter γ controls the time which the oscillator spends in these two phases, β control the gradient of the sigmoid. The coupling term S_i provides the overall input from neighboring oscillators in the network:

$$S_i = S_i^a - W_z H(z - \theta_z) \qquad (2)$$

S_i^a is the total coupling from the adjacent active neighbors of oscillator to accomplish the binding problem . The original is defined summation in Eq.(3).

$$S_i^a = \sum_{k \in N(i)} W_{ik} H(x_{ik}) \qquad (3)$$

Where W_{ik} defines the dynamic connection weight from oscillator k to i and $N(i)$ represents a set of oscillators that comprises the neighborhood of it. H stands for the Heaviside step function.

θ_z is a threshold, and W_z is the weight of inhibition from the global inhibitor, whose activity is governed by the equation:

$$\dot{z} = \phi(\sigma_\infty - z) \qquad (4)$$

Where $\sigma_\infty =1$ if $x_i > \theta_z$ for at least one oscillator i, and $\sigma_\infty =0$ otherwise.

Thus, this segmentation process is the emergent behavior of the oscillator network. For image segmentation, the LEGION network generally has two-dimensional (2-D) architecture. Each oscillator corresponds to a pixel in the given image and is connected to its eight nearest neighbors except for at the boundaries where there is no wrap around. The global inhibitor is connected to all the oscillators on the 2-D grid. It receives excitation from each oscillator and in turn exerts inhibition to each oscillator.

2.2 Extended LEGION Segmentation for Building Extraction from DSM

Due to the large number of pixels in DSM raster data, numerical integration of hundreds of thousands of differential equations of original algorithm is prohibitively expensive. Thus, an extended simplified LEGION framework is proposed. According to the purpose of building extraction from DSM, the feature detector associated with each oscillator estimates the elevation of terrain at its corresponding pixel location. Given the LEGION dynamics, the main task is to establish lateral connections based on a similarity measure. Fig.(1) shows the flow chart of extended LEGION segment for building extraction from DSM. At the beginning, cells i corresponding to pixels are initialized into a non-excitated state. Then coupling weights W_{ik} are calculated between the eight cells k adjacent to the cells i, which is based on the similarity measure. W_{ik} is represented by the following equation:

$$W_{ik} = W_{\max} / (1 + |Dissimlarity(i,k)|), k \in N(i) \qquad (5)$$

Where Dissimilarity *(i,k)* indicate the distance between pixel i and k and W_{max} indicates the maximum value of the pixels' elevation dissimilarity. Here, we use the maximum value of dissimilarities for W_{max}.

Next step is to distinguish a major oscillator block to address the fragmentation problem. Usually a lateral potential for each oscillator is applied. However, this method it is hard for LEGION to extract a building directly by segmentation, because high dense trees may contain leaders to participate in segmentation. Gray Level Co-occurrence Matrix (GLCM) homogeneity, a feature of DSM height texture, is proposed to distinguish between buildings and tall trees and locate major oscillators in building areas. One pixel windows size is used for GLCM calculation and GLCM homogeneity is represented in Eq.(6). Homogeneity returns a value that measures the closeness of the distribution of elements, is chosen to weight the value decreasing exponentially according to their distance to the diagonal. Any homogeneity values which are close to 1 are taken as leaders of LEGION segmentation.

$$H = \sum_i \sum_j \left(\left(\frac{1}{1 + (i - j)^2} \right) \frac{p(n)}{\sum_i \sum_j p(n)} \right) \qquad (6)$$

Where $p(n)$ is the DN value of pixel, and i, j is the number of rows and columns.

According to temporal oscillator correlation, the global inhibitor acts as a "metronome", which establishes a single frequency of oscillation for all objects independently of their actual input. W_z the weight of the global inhibitor plays a significant role in segmenting pixels into different groups. Yet the value of W_z usually determines by experience. In this paper, we find that there exists a relationship between DSM complexity and W_z, which helps to do the determination of W_z.

For DSM complexity, GLCM contrast attributes are applied to describe whether gray level distribution is centralized or decentralized, as contrast returns a measure of the intensity of contrast between a pixel and its neighbor over the whole image. The measurement of target occurrence is proposed to show the complexity of a target and background feature distributions as well. Target occurrence (R) is defined as Eq.(7), which is based on analysis of edge level percentages within the image(Mario et al.,2005). And the inhibition weight Wz is calculated in Eq.(8).

$$R = P_{edge} / (M \times N) \tag{7}$$

$$W_z = W_{max} \times (\text{Target Occurrence} + \text{Texture Contrast}) \times \text{Target Occurrence} \tag{8}$$

Then, one leader cell yet to be excited is selected as a self-excitable cell. The selected cell is put into the excitation state, the excitable. Cells are selected based on the coupling weights between the adjacent cells through coupling term S_i^a , and the selected cells are put into the excitation state. Here, S_i^a is applied by logarithmic operation, which was presented by Chen et al.(2000),which is shown in Eq.(9).At the same time, global inhibitory takes action to inhibit excited oscillators. Thus, based on Terman-Wang's oscillator correlation theory, oscillators for the same objects can be synchronized, while global inhibitory is used to discriminate different objects through de-synchronization. These operations of extended LEGION are repeated until no excitable cells are detected. If no excitable cell is detected, inhibition processing is performed, thereby completing the image segmentation of one region. These operations are repeated until there is no non-excited and non-inhibited leader cell any more, thereby pinpointing regions belonging to the same category from an input image and identifying them as an image segmentation regions.

$$S_i^a = \sum_{k \in N(i,1)} H(x_k) \times W_{ik} / \log\Big(\sum_{k \in N(i,1)} H(x_k) + 1 \Big) \tag{9}$$

After segmentation, morphological cleaning procedures, such as morphological opening and morphological reconstruction, are applied to the binary building segmentation images to remove small objects and to retrieve the building boundaries that are smoothed out as a result of the opening operation. Then the building boundary of each region is extracted from the detected building regions, which was measured by tracing boundary contours in a binary image mentioned by Ren, et al.(2002). Since buildings are regulated objects, solution of least squares with perpendicularity constraints is put forwarded to determine a regularized rectilinear building boundaries. Firstly, Douglas-Peucker-Algorithm gets the feature points through reducing the number of points of the original tracing point set by recursively eliminating points that fall below the threshold of a potential remaining line. In Douglas-Peucker method, the threshold distance ε affected the feature point extraction directly. Thus, threshold distance ε is defined by perimeter area ratio. Secondly, the determination of perpendicular direction is executed for regularizing the boundary. We used cosine value of near neighborhood pair feature points to judge perpendicular direction point. If the cosine value of pair points was less than 0.5, then it was considered as perpendicular direction point. Finally, all boundary points were applied in a least squares solution with perpendicularity constraints to determine a regularized rectilinear boundary. Thus, the polygons are divided into two groups according to the possibility for perpendicularity of inner angles of consecutive polygons, and then the adjustment is performed for each group.

3. Experiment and Results

3.1 Data Description

The test area of Vaihingen in Germany was covered by altogether 10 strips captured with a Leica ALS50 system, which was provided by DGPF contains Airborne Laser scanner (ALS) data（Cramer, 2010）. Inside an individual strip the average point density is 4 points/m2. The experimental area is A2 region which is characterized by a few high-rising residential buildings that are surrounded by trees. Another test area is a suburban area of scenic place in China. DSM of of A2 data was interpolated from the ALS point cloud with a grid width of 25 cm using all return information. Another experimental DSM data was interpolated with a grid width of 30 cm..They were generated by using an interpolation method nearest neighbor (NN) searches method. When interpolating in two dimensional space, the especial Quadtree is equal to the general KD-tree. Thus, space-partitioning data structure KD-tree was applied for NN search. However, there are some missing points in raw LiDAR data of urban area. There may be several reasons causing the missing points. According to the experimental data of A2 region, LiDAR data gap exists in significant changes in the ground target's height. The missing points are found manually. And the missing points on the ground are interpolated by neighbor ground points. Fig.(2) showed the DSM of experimental data.

3.2 Experiment

GLCM homogeneity was applied to distinguish buildings and tall trees and locate major oscillators in building areas, which is shown in Fig.(3).Target occurrence and texture contrast, the parameters of DSM imagery complexity, were used to define the value of the global inhibitor Wz to segment pixels into different groups. Tab.(1) showed the result of DSM complexity measurements and the weight of inhibition. Thus the extended LEGION scheme was applied to extract buildings and remove trees from DSM segmentation in complex urban or suburban areas.

Table. 1 DSM complexity measurements and the weight of inhibition

	Target Occurrence	Texture Contrast	Real Wz	Estimated Wz
A2	0.3358	0.1116	38	38.31
Suburban	0.3228	0.8549	31	31.33

Morphological opening and morphological reconstruction were applied to the binary building images in order to remove small objects after segmentation. Here we used square structuring elements in 5 pixels width for A2 area and 3pixels for suburban area to construct the morphological operation. The results were shown in Fig.(4).

The Douglas-Peucker-Algorithm was used to get the feature points of the original point set by recursively eliminating points. Then the solution of least squares with perpendicularity was applied to determine regularized building boundaries. Finally the building outlines were converted in DXF format for ISPRS test evaluation. Fig.(5) showed the regularized building boundaries.

3.3 Result Analysis

Our methods detected 8 buildings in Vaihingen A2 area and 6 buildings in suburban area. Pixel-based evaluation and object-based evaluation on Point-in-Polygon were used as evaluation based on the method described in Rutzinger et al., (2009). The evaluation of Vaihingen A2 area was provided by Rottensteiner et al. as result of ISPRS Benchmark on Urban Object Classificaton and 3D Building Reconstruction, while the evaluation of suburban area was calculated manually by ourselves.

In pixel-based evaluation, completeness represented Producer's Accuracy, correctness represented User's Accuracy, and quality represented balances of completeness and correctness. Fig.(6) showed the pixels classified as True Positive (TP), False Positive (FP) and False Negative (FN) in the pixel-based evaluation, which indicated that the algorithm had detected the majority of the buildings, without too many FPs performance. Tab.(2) showed the result of pixel-based evaluation.

Table 2 Results of a pixel-based evaluation

	Pixel completeness	Pixel correctness	Pixel quality
A2	88.5%	98.9%	87.6%
Suburban	87.9%	98.3%	86.6%

Since object-based evaluation techniques are less sensitive to errors at the building outlines and can be related to building parameters (Rottensteiner et al., 2005), object-based evaluation based on Point-in-Polygon Tests was also used to evaluate. Tab.(3) showed the results of the performance evaluation based on the PIP test.

Table 3 Results of an object-based evaluation using PIP test

	Object completeness	Object correctness	Object quality
A2	71.4%	100.0%	71.4%
Suburban	70.3%	100.0%	70.3%

The results of evaluation indicated that building extraction from DSM in a complex urban area by using extended LEGION can achieve considerable accuracy.

4. Conclusion

The study presented in this paper has used the neural oscillator network LEGION approach and applied it to building extraction from DSM. The extended LEGION method needs no assumption about the underlying structures in DSM data and no prior knowledge regarding the number of regions. This method successfully segmented real DSM data, which shows that it may represent a generic DSM segmentation method. Building objects in urban and suburban areas DSM can be detected efficiently and effectively. The boundary regularization method takes rectangularity constraints and arc constraints into account, and thus produces promising results.

Acknowledgements

This work is supported by National Basic Research Program of China (2012CB957702) and the Innovation Program of Shanghai Municipal Education Commission (No. 11YZ290) .We acknowledge the data set is a subset of the data used for the test of digital aerial cameras carried out by the German Association of Photogrammetry and Remote Sensing (DGPF) ,which referenced from M. Cramer. And we would like to thank for The ISPRS Benchmark on Urban Object Classification and 3D Building Reconstruction, which provided by F. Rottensteiner, G. Sohn, M. Gerke, C. Baillard, S. Benitez, U. Breitkopf.

References

Al-Harthy, A., and Bethel J., 2002. Heuristic filtering and 3D feature extraction, Proceedings of ISPRS Commission III Symposium, Graz, Austria, unpaginated CD-ROM.

Chen K., Wang D L., Liu X W., 2000.Weight Adaptation and Oscillatory Correlation for Image Segmentation.IEEE Transactions on Neural Networks,11(5):1106-1123

Cramer, M., 2010. The DGPF test on digital aerial camera evaluation – overview and test design. Photogrammetrie Fernerkundung Geoinformation, 2:73-82.

Li, Y., Zhu, L., et al. 2010. A refined marker controlled watershed for building extraction from DSM and imagery. International Journal of Remote Sensing 31(6), 1441-1452.

Lu, Y.H., Trinder, J.C., et al. 2006. Automatic building detection using the Dempster-Shafer algorithm. Photogrammetric Engineering and Remote Sensing 72 (4), 395-403.

Mario, I., Chacon, M., et al. 2005. Image complexity measure : A human criterion free app roach .Proceedings of IEEE Annual Meeting of the North American Fuzzy Information Processing Society. Detroit, 28 June, 2412246.

Miliaresis, G. and Kokkas, N., 2007. Segmentation and object-based classification for the extraction of the building class from LIDAR DEMs. Computers & Geosciences, 33(8): 1076–1087.

Ren, M.W., Yang,J.Y., Sun, H.,2002. Tracing boundary contours in a binary image, Image and Vision Computing, 20(2),125-131

Rottensteiner, F., Trinder, J., et al.2005. Using the Dempster-Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. Information Fusion 6(4),283-300.

Rutzinger, M., Rottensteiner, F., et al. 2009. A comparison of evaluation techniques for building extraction from airborne laser scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2(1),11-20.

Tullis, J. A., Jensen, J.R., 2003. Expert system house detection in high spatial resolution imagery using size, shape, and context. Geocarto International, 18(1), 5-15.

Walter, V., 2004. Object-based classification of remote sensing data for change detection. ISPRS Journal of Photogrammetry and Remote Sensing 58(3-4), 225-238.

Wang, D.L., Terman, D.,1995. Locally excitatory globally inhibitory oscillator networks. IEEE Transactionson Neural Networks, 6, 283-286.

Wang, D.L., Terman, D.,1997. Image segmentation based on oscillatory correlation. Neural Computation 9, 805-836.

Wang, D.L. 2007 .Computational scene analysis. In Challenges for Computational Intelligence, W. Duch and J. Mandziuk, Ed, Springer, Berlin.

Wang, Z., Schenk T., 2000. Building extraction and reconstruction from LiDAR data, International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B3. Amsterdam, The Netherlands, unpaginated CD-ROM.

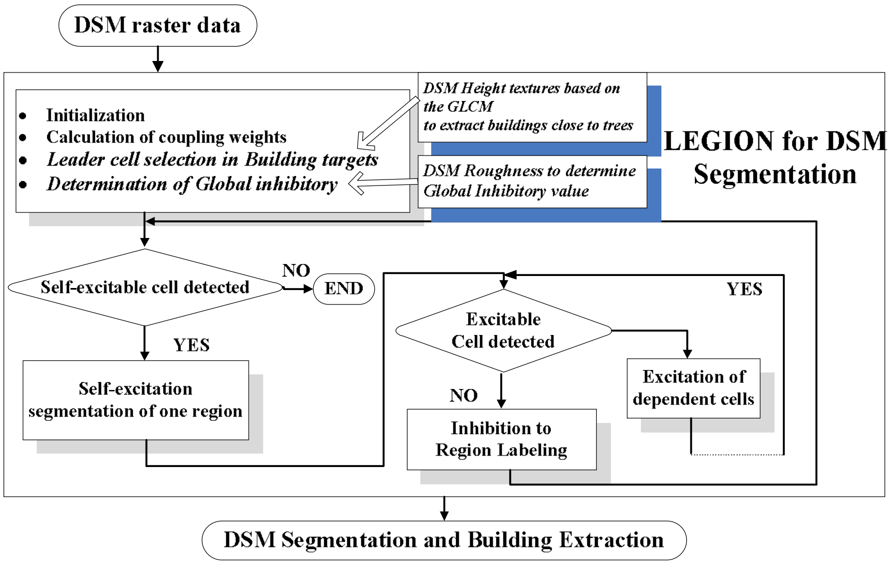
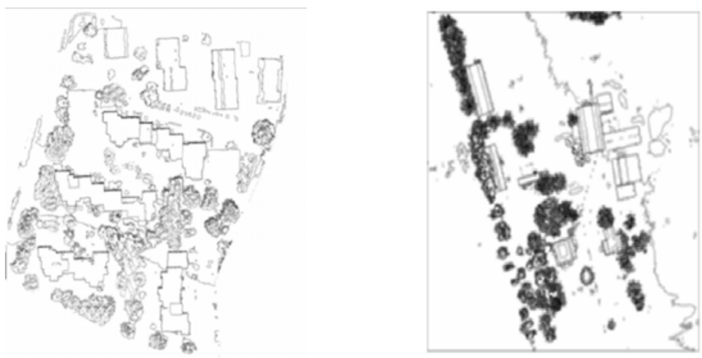


Figure. 1 Flow chart of Extended LEGION Segment for Building Extraction from DSM

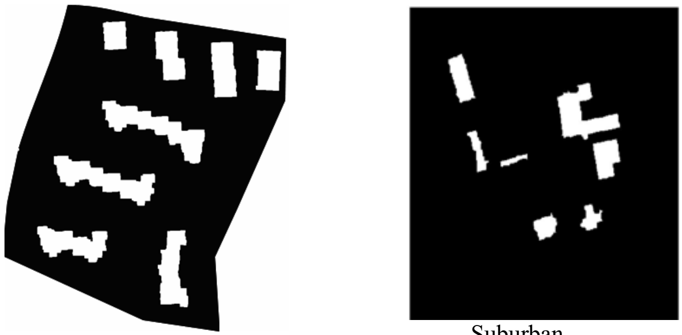
A2 Suburban

Figure 2. DSM of experimental data



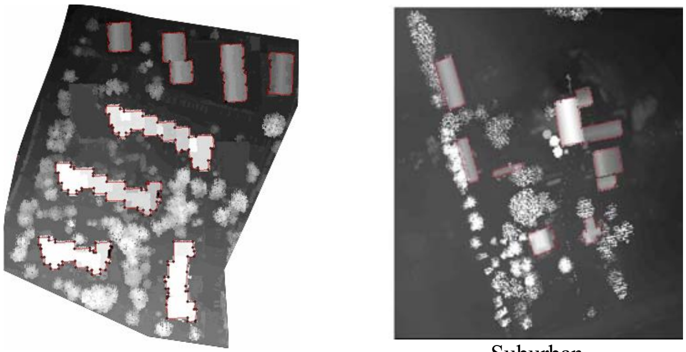
A2 Suburban

Figure 3 GLCM homogeneity result of experiment data(homogeneity in light)



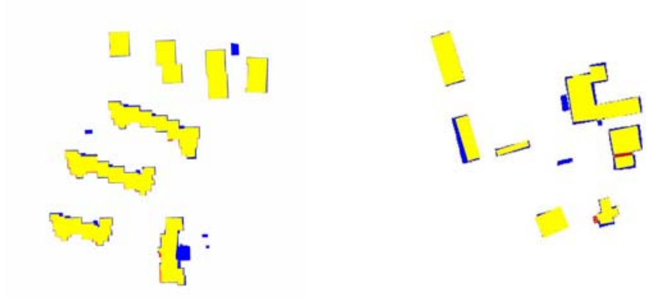
A2 Suburban

Figure 4.Building Segmentation Results



A2 Suburban

Figure 5 Regularized building boundaries of experimental data



A2 Suburban

Figure 6 Pixel-based evaluation of the building detection (Yellow: TPs; Red: FPs; Blue: FNs)