A Watershed-Canny Based Approach for Building Footprint Extraction from Very High Resolution Optical Image

Faycal SAIDI[a,*] and Jie CHEN[b]

School of Electronics and Information Engineering,
Beihang University,37 Xueyuan Road, Haidian District, Beijing 100191, China
[a]faysaidi09@gmail.com, [b]chenjie@buaa.edu.cn
*Corresponding author

Keywords: Building Footprint Extraction, VHR Optical Image, Canny Edge Detector, Watershed Segmentation.

Abstract. The advanced very high resolution (VHR) sensors are capable of achieving sub-meter resolution, which offers the opportunity for a fine level of analysis of man-made structures. In this paper, we present a method for the extraction of 2-D building footprints from VHR optical scenes. The data sets include Worldview-2 (0.5m) image and cover urban area of San Francisco. Our main idea was to combine edge based detector, regions based segmentation and non-building masks. In the first step, Canny operator was used to extract edge from the optical image and morphological operations were used to remove small edge. In the second step, the Watershed transform was used for segmentation of optical image and morphological operations were used to remove small regions. In the third step, we computed two non-building masks which are vegetation mask and shadow mask. These two masks were applied first to filter Watershed segmentation result. In the fourth step, we have translated the Canny contour image in both directions and computed the Correlation Coefficient between Canny contour and filtered Watershed contour for each value of translation, and we got the best translation parameters when Correlation Coefficient was maximal. In the last step, the matched Canny edge image was combined separately with: 1) vegetation mask, 2) shadow mask and 3) filtered Watershed segmentation. The three results were then combined (logical "or") to obtain the final building footprint. The obtained results demonstrated that our approach performs well and improves the boundaries extraction based on Watershed segmentation.

1. Introduction

In the last decade, the development of remote sensing technologies and the launch of high-resolution remote sensing satellites have increased the amount of available metrics images [1]. Very high resolution satellite images, which are an important source of information for decision makers, are used for many tasks like change detection, image fusion [2], urban growth monitoring and damage assessment.

Extraction of the land use information from remote sensed data is a very important task in the Geographic Information Systems (GIS). One of the important land types is building region. The extracted building regions are useful for mapping and map updating, disaster monitoring, as well as other applications. Traditionally, building regions are extracted from image data [3].

Considering the migration of thousands of people from rural to urban areas, the land cover classes in urban and suburban areas are changing rapidly, and this trend will increase in the future with urbanization. Also, and without any human intervention, Earth's surface is highly dynamic itself, it is changing continuously. Thus, continuous monitoring of changing Earth is very important and extremely useful [4].

Thereby, building detection using remote sensing images is one of the most challenging problems of target detection. For the automated building detection system, various algorithms have been developed for a better automatic building extraction but they do not provide an exact solution. In literature, the methods aim to solve this problem with a great variety of approaches. These approaches can be grouped under two methods: supervised and unsupervised methods. The unsupervised approaches, detect the buildings using predefined rule-based models and unsupervised classifiers. One of the most popular approaches in this group is the use of shape based features which represent the rectangular structure of the roofs [5]. Some algorithms employ techniques to discriminate and remove the irrelevant regions from the image, then focus on the regions which include buildings. For example, authors in [6] formulated the urban-region and building detection problems using graphical models. In [7], they used the shadow evidence to focus on building regions which models the spatial relationship between buildings and their shadows [8].

2. Research Frame Work and Theoretical Background

The increasing availability of very high resolution (VHR) images from satellites like GeoEye-1 and Worldview-1/4, provides great opportunity for extracting buildings from satellite images. Edge detection and segmentation are fundamental tasks in image analysis. However, conventional algorithms often miss parts of the true boundaries. Considering that building areas and their surroundings are presented with various intensity values and complex features, obtaining the segmentation of building footprints from satellite images is a complex process. In general, region-based segmentation methods can be used to establish the boundaries of building structures. These methods divide the image into similar and homogenous regions. However, using these approaches for building detection in several cases where spectral heterogeneity exists, over-detection or under-detection is usually noticed.

In this paper, a combined building footprint extraction approach based on Watershed segmentation and Canny edge detector is proposed. In this approach, we use Canny to extract boundaries and we use Watershed transform to segment the image. We also use vegetation and shadow masks to remove non-building area and then, we combine the results to obtain the refined building footprint.

2.1. Discrete Canny Algorithm

The main steps of discrete Canny algorithm are as follows [9]:
- Smooth the original image in order to reduce the noise;
- Convolve the image with edge detection operator, like Sobel, to get the intensity gradient of the image:

$$G=\sqrt{(G_x^2+G_y^2)} \; ; \theta=\arctan(G_x/G_y) \qquad (1)$$

where $G_x{}^2$ and $G_y{}^2$ are the gradient on x and y axis, G is the gradient on the maximum direction and θ is the angle between maximum gradient direction and x axis;

- Suppression of Non-maximum: the edges are set to a width of one pixel. Every pixel which belongs to an edge doesn't have a neighbor pixel, except in edge direction.
- Compute a hysteresis threshold th_1 and th_2 with $th_1 > th_2$: if the grey value of a pixel is larger than th_1, it is considered as edge. If it is smaller than th_2, it is not. For the grey values between th_1 and th_2, if the grey values of its abutting pixels are larger than th_1, they are edges, otherwise they are not.

2.2. Watershed Transform

In order to divide the panchromatic image into spectrally homogeneous regions, image segmentation was performed using the Watershed segmentation algorithm. Considering the objectives of our work, this algorithm seems to be adapted and provides satisfactory results. Also, the use of the Watershed algorithm allows the placement of boundaries at the significant edges and the produced regions are closed and connected, whereas edge-based techniques usually lead to disconnected boundaries.

The Watershed segmentation technique is based on a simple heuristic approach that consists in analyzing the gray level of the image pixels in an ascending order and performing region growing[10]. Because of the regions in the image characterized by small variations in gray levels, in practice, the Watershed segmentation is often applied to the gradient of an image rather than to the image itself [11].The Watershed segmentation technique uses, in addition to the gradient image, a seed image calculated from the gradient image. By gathering into the region, the pixels that are the closest to the corresponding seed, the growing process determines the region associated with each seed. This will satisfy a certain homogeneity in the gray level [10].

However, the obtained image is generally over-segmented. So, in order to overcome these shortcomings, we propose to combine the Watershed segmentation with Canny edge detector and some morphological filters.

2.3. Non-Building Mask

We also produce a vegetation mask image which is obtained by thresholding the Normalized Difference Vegetation Index (NDVI) of the multispectral image and the Shadow Index (SI) mask image. The NDVI and SI are given by [12]:

NDVI=(NIR-RED)/(NIR+RED)   (2)

SI= (RED+GREEN+BLUE+3NIR)/6   (3)

2.4. Correlation Coefficient

The Correlation Coefficient (CC) is a number representing the similarity between two images based on their respective pixels intensities. Karl Pearson defined the Pearson product-moment correlation coefficient r. It was the first formal correlation measure and is widely used in statistical analysis, pattern recognition and image processing. The Pearson's Correlation Coefficient is defined as:

$$r = (\textstyle\sum_i \sum_j (x_{ij}-x_m)(y_{ij}-y_m))/ \sqrt{((\textstyle\sum_i \sum_j (x_{ij}-x_m)^2)(\textstyle\sum_i \sum_j (x_{ij}-x_m)^2))} \qquad (4)$$

Here, x_{ij} and y_{ij} are the images to compare, whereas the subscript indices i and j refer to the pixel location in the image. Also, x_m and y_m are mean intensity values of 1st and 2nd image respectively.

The Correlation Coefficient has the value $r = 1$ if the two images are absolutely identical, $r = 0$ if they are completely uncorrelated and $r = -1$ if they are completely anti-correlated [13].

3. Proposed Approach and Experimental Results

3.1. Methodology

For the extraction of a 2-D building footprint from the optical image, we propose to combine two methods. First, we use Canny edge detector to extract features from the optical image. Second, we use the Watershed transform for the image segmentation [14]. Otherwise, in order to remove non-building class, we produce a vegetation index (NDVI) mask and a shadow index (SI) mask. These two masks were applied first to filter Watershed segmentation result. Then, we matched the features by using Correlation Coefficient. We translated the Canny contour image in both directions and we computed the CC between Canny contour and filtered Watershed contour for each value of translation, and we got the best translation parameters when Correlation Coefficient was maximal. In the last step, the matched Canny edge image was combined separately with: 1) vegetation mask, 2) shadow mask and 3) filtered Watershed segmentation. The three results were then combined (logical "or") to obtain the final building footprint. The Figure 1 shows flowchart of the proposed method.
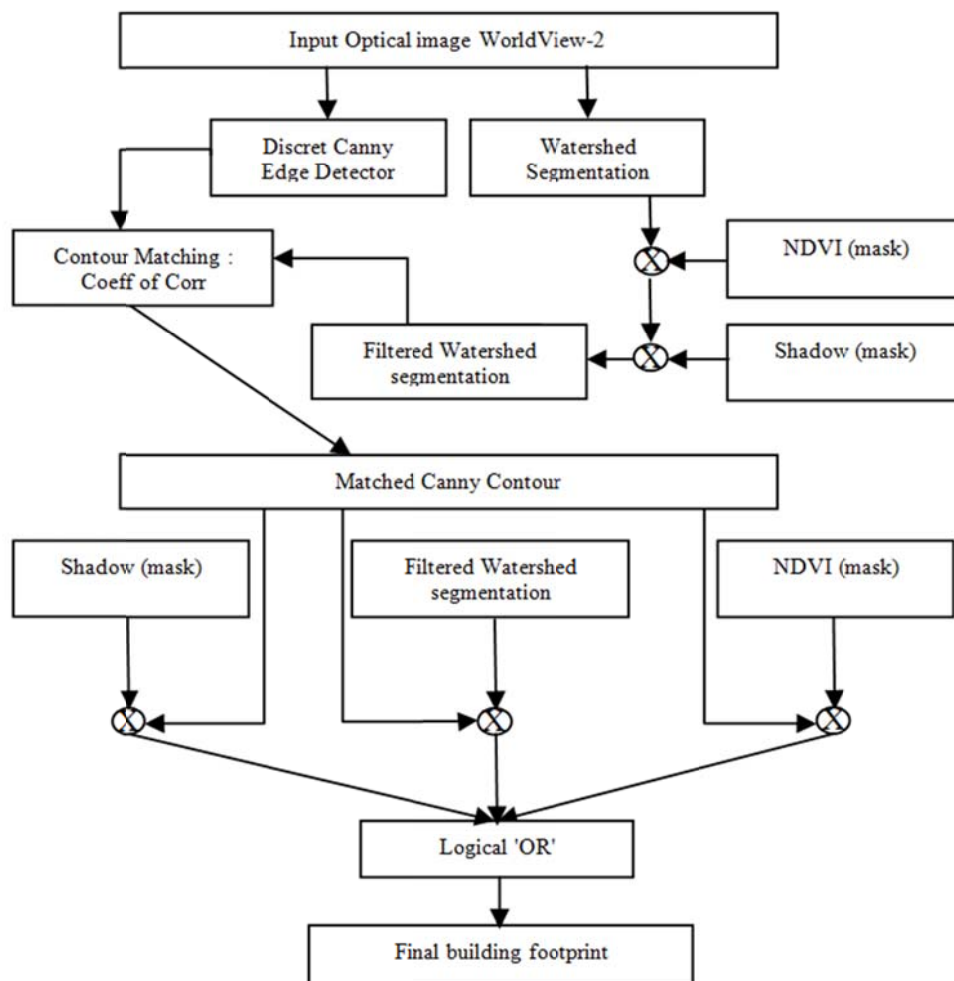

Figure 1 Flowchart of proposed method.

3.2. Dataset

In this work, the proposed approach was tested using two very high spatial resolution WorldView-2 images (0.5m), acquired on October 9th 2011, covering a high density urban region of San Francisco. The images are projected in Universal Transverse Mercator (UTM) coordinate system, Zone 10. The dimensions of images are (1879 x 1350 and 1120 x 672 pixels). The Figure 2 shows the input optical images.
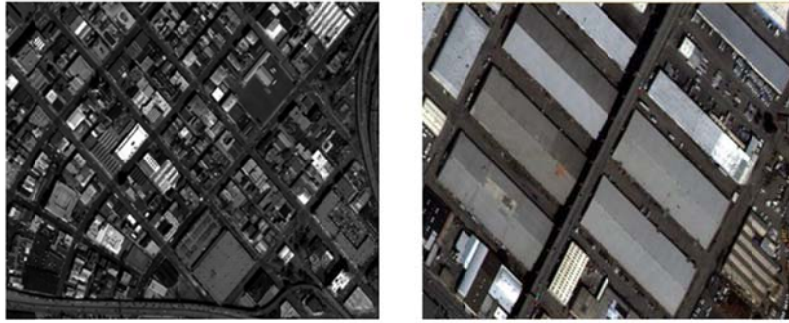

Figure 2 Original input WorldView2 (0.5m) optical images San Francisco 1&2.

3.3. Simulation

In this section, we will perform the proposed approach for building footprint extraction which consists in the following steps:

3.3.1. Application of Watershed Segmentation

We used Watershed segmentation on the grey scale of optical image and then a morphological filter was applied to remove small regions. Here, we extracted objects from the binary image using features properties which is area. Therefore, Figure 3 shows the results of Watershed segmentation.
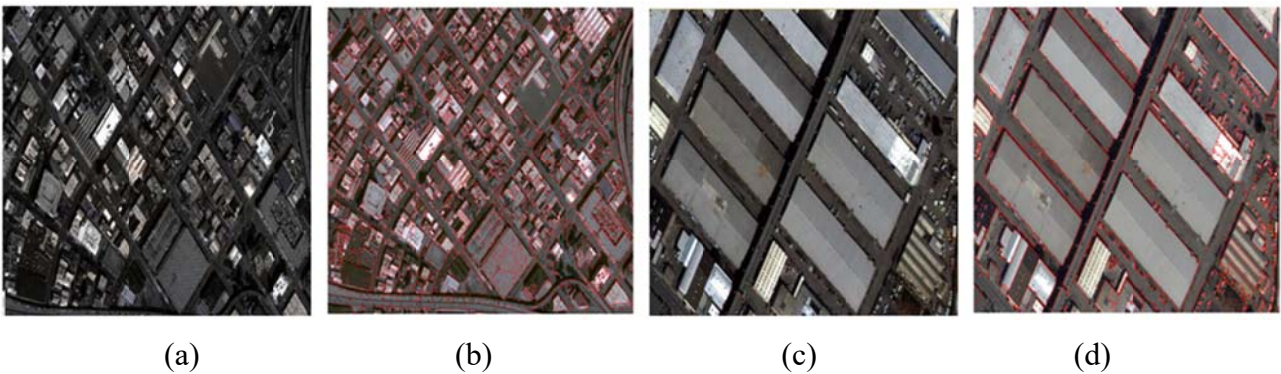

| (a) | (b) | (c) | (d) |

Figure 3 Watershed segmentation, (a) Original San Francisco 1 image, (b) Watershed segmentation boundaries, (c) Original San Francisco 2 image, (d) Watershed segmentation boundaries.

3.3.2. Application of the Canny Edge Detector

In this step, contours in the Worldview-2 optical image were extracted using the discrete Canny edge detector (with σ =√2). Then we used a morphological filter to remove small edge. The Figure 4 illustrates the contours detected from the optical image.
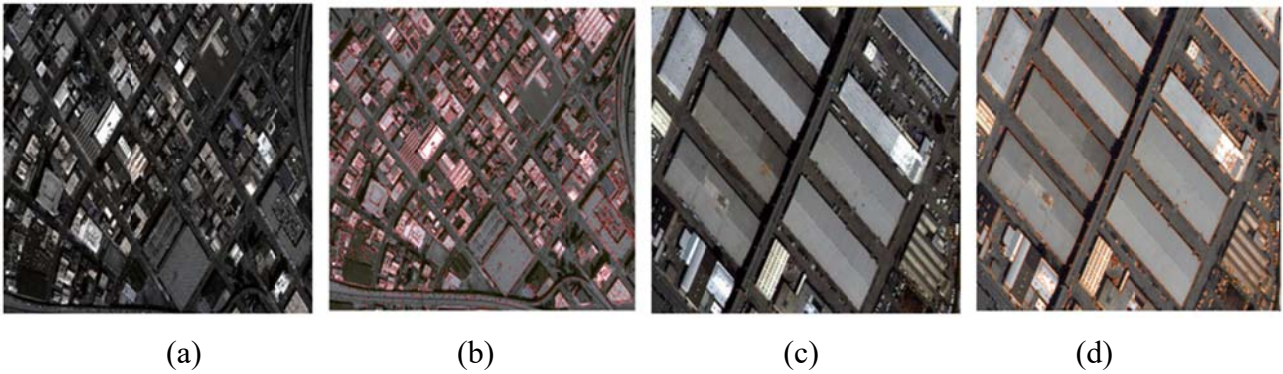
Figure 4 Contours detected by Canny, (a) Original San Francisco 1 image, (b) Canny detector boundaries, (c) Original San Francisco 2 image, (d) Canny detector boundaries.

3.3.3. Computation of Non-Building Mask

In this step, we computed a vegetation mask by thresholding the normalized difference vegetation index (NDVI) given by equation 2 and, a building shadow mask by thresholding the shadow index (SI) given by equation 3. Then we used a morphological filter to smooth the results. The Figure 5 illustrates the final vegetation and shadow masks.
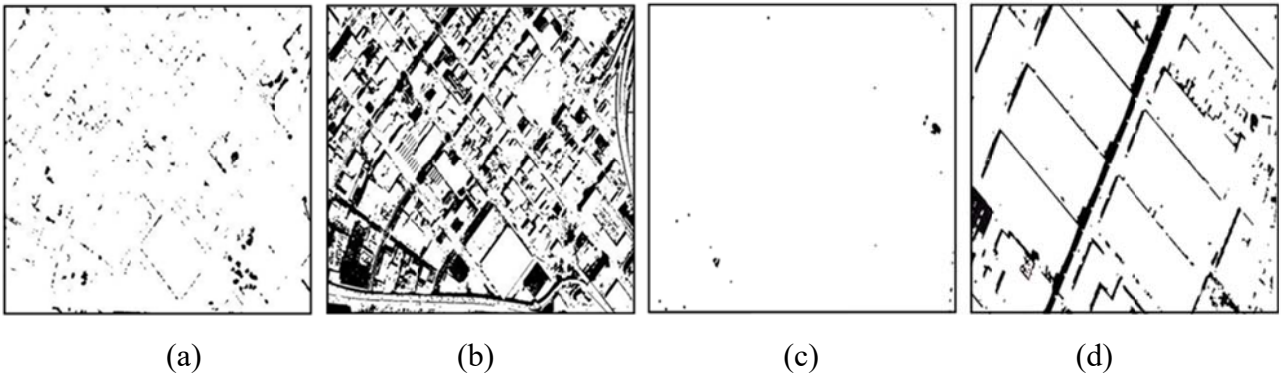

Figure 5 (a) vegetation mask, (b) shadow mask for San Francisco1; (c) vegetation mask, (d) shadow mask for San Francisco2.

3.3.4. Edge Matching and Combination

In this step, we first applied the non-building masks to the Watershed segmentation then, we computed the Correlation Coefficient (CC) between Canny contour and Watershed contour. We translated the Canny contour image in the range of *-1* to *1* pixels (*step=0.1*) in both directions and we computed the correlation coefficient and then, we found the best values of (*x_shift, y_shift*) corresponding to the maximal value of CC. In our case, we got the best result for the following transformation parameters: *x_shift = -0.1* pixel, *y_shift = -0.4* pixel.

The matched edge image was combined separately with: 1) vegetation mask, 2) shadow mask and 3) Watershed segmentation. The three results were then combined (logical "or") to obtain the final building footprint. The Figure 6 illustrates the final footprint extracted from optical image.
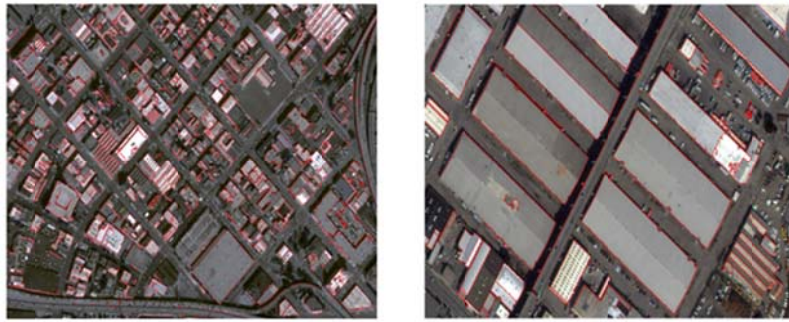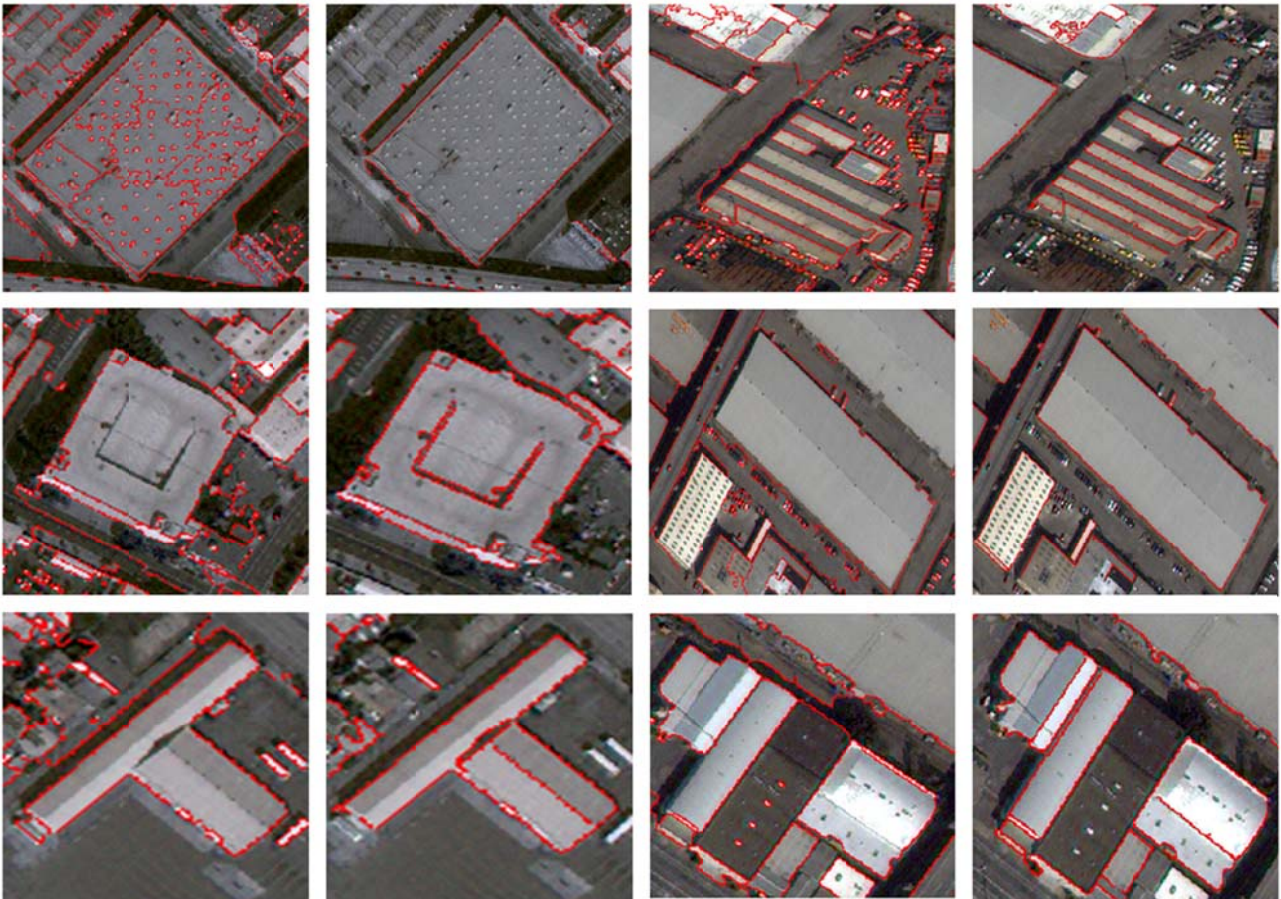
(a)   (b)

Figure 6 Building footprint, (a) final building footprint of San Francisco1, (b) final building footprint of San Francisco2.

3.3.5. Evaluation

In order to evaluate the performance of the proposed approach, it has been tested on two sub-scenes. The first one is very dense urban area and the second one contains hangars, so we propose to use:

(a)   (b)   (c)   (d)

Figure 7 Evaluation of the proposed approach (a) Watershed segmentation contours of San Francisco1, (b) proposed approach contours of San Francisco1, (c) Watershed segmentation contours of San Francisco2, (d) proposed approach contours of San Francisco2.

- Qualitative (visual) evaluation: a visual evaluation was done by comparing the result of the proposed approach with Watershed segmentation approach. The Figure 7 shows that the proposed method eliminates the over detection caused by Watershed segmentation in both datasets.
- Quantitative evaluation: a quantitative evaluation was also done for the second dataset. Indeed, from 23 building identified in the dataset, 17 were correctly and completely detected, 4 building were correctly but partially detected and 2 missed building.

4. Conclusion

In this work, we described an approach for detecting buildings using very high spatial resolution optical image. After an initial over segmentation based on Watershed segmentation, we used the Canny edge detector and non-building masks, which are vegetation mask and shadow mask, to improve the detection results. We tested the proposed approach on two scenes with different building characteristics. The tests showed that the proposed approach is able to detect buildings with different colors and shapes. The evaluation was done visually, by comparing the results of our approach with the results of Watershed segmentation, and quantitatively by calculating the number of detected buildings compared to the number of originally identified buildings. The evaluation confirmed that the proposed approach performs well and improves the building footprint extraction based on Watershed segmentation.

Acknowledgments

References

[1] Suri, S. and Reinartz, P. (2010) Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas. IEEE Transactions on Geosciences and Remote Sensing, vol.48. no. 2, 939 - 949, doi : 10.1109/TGRS.2009.2034842.

[2] Chen, H.M., Varshney, P.K. and Arora, M.K. (2003) Mutual information based image registration of remote sensing data. International Journal of Remote Sensing, 24(18), 3701-3706, doi:10.1080/0143116031000117047.

[3] Teo, T.A. and Chen, L.C. (2004) Object-based building detection from Lidar data and high resolution satellite imagery. 25thAsian Conference on Remote Sensing.

[4] Mishra, B. and Susaki, J. (2014) SAR and optical data fusion for land use and cover change detection. IEEE International Geosciences and Remote Sensing Symposium, doi:10.1109/IGARSS.2014.6947540.

[5] Saeedi, P. and Zwick, H. (2008) Automatic building detection in aerial and satellite images. IEEE International Conference on Control, Automation, Robotics and Vision, vol. 1,623–629.

[6] Sirmacek, B. and Unsalan, C. (2009) Urban-area and building detection using sift keypoints and graph theory. IEEE Transactions on Geosciences and Remote Sensing, vol. 47, no. 4,1156–1167.

[7] Ok, A.O., Senaras, C. and Yuksel, B. (2013) Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery. IEEE Transactions on Geosciences and Remote Sensing, vol. 51, no. 3,1701–1717.

[8] Senaras, C., Ozay, M. and Vural, F.T.Y. (2013) Building detection with decision fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 6, no. 3.

[9] Türmer, S. (2009) Automatic Registration of High Resolution SAR and Optical Satellite Imagery in Urban Areas. Thesis.

[10] El Merabet, Y., Meurie, C., Ruichek, Y., Sbihi, A. and Touahni, R. (2015) Building Roof Segmentation from Aerial Images Using a Line-and Region-Based Watershed Segmentation Technique. Sensors, 15, 3172-3203, doi:10.3390/s150203172.

[11] Salman, N. (2006) Image Segmentation Based on Watershed and Edge Detection Techniques. The International Arab Journal of Information Technology, vol. 3, no. 2.

[12] Bouziani, M., Goïta, K. and He, D.C. (2007) Change detection of buildings in urban environment from high spatial resolution satellite images using existing cartographic data and prior knowledge. IEEE International Geosciences and Remote Sensing Symposium, 2581-2584.

[13] Kaur, A., Kaur, L. and Gupta, S. (2012) Image Recognition using Coefficient of Correlation and Structural Similarity Index in uncontrolled environment. International Journal of Computer Applications, vol. 59, no. 5,0975 – 8887.

[14] Yan, L. and Li, P. (2014) Building footprints extraction methods based on marker-controlled Watershed segmentation and local Radon transformation. Information Technology Journal, 13: 1903-1907.