Detection of Residential Buildings to Estimate Population in Lebanon using GeoEye Images

Kamel Allaw, Jocelyne Adjizian Gerard, Makram Chehayeb, Nada Badaro Saliba

1CREEMO, Geography Department, Saint Joseph University, Beirut, Lebanon.
2Surveying Department, Islamic University of Lebanon, Lebanon.

Correspondence should be addressed to Makram Chehayeb, makram.ch92@gmail.com

Publication Date: 18 July 2019

DOI: https://doi.org/10.23953/cloud.ijarsg.419

Abstract Scholars in urban planning and Geography are increasingly interested in grasping demographic information using Remote Sensing data. The accurate detection of residential buildings from satellite images seems to be essential in this domain. This paper has a dual purpose: It aims firstly at developing an automatized method for residential buildings extraction, then, evaluating the relationship between residential building characteristics (number, area, and volume) and demographic data. To do so, a dual phasic methodology is proposed. During the first phase, the extraction of residential buildings has been done using a transformation into HSI representation where the buildings correspond to the higher values of band I. After that, the image has been transformed into vector and the forms of the buildings have been adjusted using convex hull tool in ArcGIS. The identification of residential buildings has been done using statistical data. The volumes of buildings has been calculated using MATLAB script. During the second phase, a multivariate regression has been established and a strong relationship ($R^2=0.87$) has been found between the volume of buildings and the population data.

Keywords Building detection; Convex hull; High resolution satellite image; HIS; Population; Supervised classification

1. Introduction

Building detection is the subject of many studies in the field of Remote Sensing due to the wide range of applications that become available thanks to the generated data. In the scientific literature, several approaches were adopted which differ according to the type of adopted method and the used data. A semi-automated approach has been adopted by Rüther et al. (2002) using an active contour model and the dynamic programming optimization technique. As well, Koç et al. (2005) who developed an approach using high resolution images, by applying classification, and using the digital surface model (DSM) and object extraction techniques. An automated approach has been adopted by Guo et al. (2002) who used high resolution IKONOS satellite images and adopted a snake-based approach for 2D building outlines’ extraction from, airborne laser scanning system has been used to capture height data. Moreover, a pair of optical and synthetic aperture radar (SAR) images have been used by Tupin et al. (2003) to automatically extract building outlines. They first used SAR image to extract partially potential building footprints, then, they used the optical one to detect shapes based the extracted lines.
In addition, Jin X. and Davis C.H. (2005) have automatically extracted buildings from high resolution satellite imagery using contextual, structural, and spectral information. This approach is characterized by three different extraction strategies: A shape analysis differential morphological profile (DMP) has been used to generate and verify building hypotheses, shadows modeling to generate relevant information in order to determine the characteristics (position and size) of adjacent buildings and spectral information has been used to detect small bright buildings which cannot be detected based on contextual and structural information. The major disadvantage of this approach is its inability to detect ceiling roof buildings where only 70% of the buildings are successfully detected. Furthermore, GÜDÜCÜ (2008) proposed a method for limit extraction based on color segmentation and using high resolution satellite images. Lee et al. (2003) proposed a new method based on Supervised classification and Hough transformation and using Ikonos images to accurately detect building roofs. Koç D. and Türker M. (2005) developed an approach using high resolution images, by applying classification, and using the digital surface model (DSM) and object extraction techniques. In addition, Inglada, J. (2007) proposed a new method for man-made objects’ detection based on support vector machines classification (SVM) and using high resolution optical remote sensing imagery. In this method, the original bands of SPOT 5 satellite images have been used for SVM learning and the extra bands (such as NDVI, nDSM and other texture measures) were used to find building patches. The additional bands increased the accuracy of the method by 10%. Salar Ghaffarian, Saman Ghaffarian (2014), applied an approach to detect buildings by collecting field data to make a supervised classification where pixels are classified into four classes, and applying morphological operations. Some methods based on classification are proposed for detecting and extracting buildings from remote sensing images this, by indicating a set of training sample to take several classes for supervised classification and using high resolution satellite images.

The previously presented methods might be accurate in structured urban areas. However, those methods are not compatible with unorganized areas where finding a pattern seems to be difficult. For that, it is essential to find a method to be accurate and useful for unorganized areas. In this context, Lebanon, as the majority of third world countries, witnessed a rapid and uncontrolled urban sprawl as a result of the corruption and the limited planning regulations (CDR, 2005). As a result, the establishment of plans to handle this situation faces huge challenges since it is difficult to obtain the required data related to constructed areas and its population data. For that, this situation imposes the elaboration of precise methods to provide researchers and specialists in urban planning with the necessary information about the built-up and population characteristics. Consequently, the aim of this study to design and test an appropriate method for building detection and population estimation.

2. Case Study

The Lebanese Republic, occupying an area of 10452 km², is located in western Asia on the eastern shore of the Mediterranean Sea between latitudes 33° and 35° N and longitudes 35° and 37° E (Figure 1). It borders the Sea in the west, Syria in the north and east, and Occupied Palestine in the south. In order to extract the necessary data for this study, a sample of 30 villages (e.g. unorganized areas) have been chosen. Those municipalities are arbitrarily chosen and geographically well distributed through the country.
3. Methodology and Data

3.1. Building Detection

To extract the buildings from the Geo-eye satellite image, three steps have been followed. First, a conversion from RGB color space to HSI color model has been done using “RGB to HIS tool” in ERDAS Imagine Software. It is important to note that the HSI color model represents each color with three components: hue (H), saturation (S), intensity (I) (Figure 2).

The Hue component (H) describes the color itself in the form of an angle between 0 and 360 degrees, the Saturation component measures the degree of purity of a color, that is, the amount of gray added to the color. The range of the S component is [0, 1]. The Intensity is the degree of brightening or darkening of a color. The range is between [0, 1] where 0 means black, 1 means white. The essential advantage of the HSI model is to distinctly separate the information of brightness, from the hue and saturation. The advantage of this model for image processing and color compositions is that it improves the visual quality of an image since.

After that, the converted HSI image can be displayed on 'ARCMAP', according to the three components H, S, and I. To determine the values corresponding to buildings, the same satellite image has been represented according to each of the three bands (H, S, I) (Figure 3). By visual interpretation, it’s easy to distinguish the color of the interval representing the buildings and the other elements which took the same I value (rocky land, roads ...). As a result, the buildings correspond to the highest values of Band I. Then, the image shown according to I band, has been converted into vector data (shape file), so that we can eliminate all elements that are non-buildings such as rocky terrain, some road portion... By this conversion, all polygons of negligible area have been eliminated (e.g. the polygons which have a surface less than 10 m²).
Figure 2: a) Satellite imagery represented Bikfaya, b) Displayed image according to the band H, c) Displayed image according to the band I, d) Displayed image according to the band S

Figure 3: Polygons represented the building in Bikfaya
3.2. Generation of Polygons Representing Buildings

In the field of geometry, convex polygons are those whose diagonals are always internal and whose internal angles do not exceed the radian pi or 180 degrees. Another way to determine if a polygon is convex is to draw segments between two points of the figure, whatever its position, if these segments are always internal, it will be a convex polygon.

'Convex Hull' is the smallest polygon that completely surrounds a set of points. This method is used to adjust the shape of the polygons that represent the buildings in all the selected municipalities. By generating these convex polygons, we can clearly distinguish all the detected buildings, and then calculate their surfaces. Because of some errors due to the existence of other objects with similar values of buildings such as roads and land, certain building areas are affected, since the polygons generated represent a larger area than the real one, which is manually adjusted as much as possible before calculating the areas of buildings (Figure 4).

![Figure 4: Application of convex Hull in a) Bikfaya, b) Nabatieh ElFaouka](image)

3.3. Identification of Residential Buildings

In order to identify the residential buildings, the area has been considered as a key identifier. In fact, the area of residential buildings in Lebanon is usually between 80 and 130 m$^2$ (Table 1). In practice, the areas of generated polygons have been divided into three categories: buildings with small areas, large areas' buildings and residential buildings. As a result, all polygons with surface less than 30 m$^2$ have been eliminated because they do not represent residential buildings. Moreover, buildings with an area between 30 and 230 m$^2$ have been considered as residential buildings because this interval include the majority of residential buildings' areas. As well, the polygons whose surfaces vary between 230 and 1000 m$^2$, have been classified as residential buildings since they may represent one or two residential buildings. As a result, polygons whose areas are greater than 1000 m$^2$ have been eliminated because they don't represent residential buildings.
Table 1: Housing distribution following the surface of the residential building in Lebanon (The central administration of statistics in Lebanon)

| Residential building area | Lebanon 2014 % |
|---------------------------|-----------------|
| Less than 30 m²           | 0.9             |
| 30 – 80 m²                | 14.1            |
| 80 – 130 m²               | 41.9            |
| 130 – 180 m²              | 21.3            |
| 180 – 230 m²              | 14.8            |
| More than 230 m²          | 5.8             |
| Unknown                   | 1.1             |
| Sum                       | 100             |

3.4. Buildings’ Areas

The areas of buildings can be calculated by multiplying the average area of buildings in each municipality by the number of buildings previously calculated. This calculation method may reduce the error resulting from the detection of buildings or residential buildings identification.

3.5. Buildings’ Volumes

The volume of buildings can be calculated by multiplying the mean area of buildings in each municipality with the number of buildings and the corresponding mean height.

\[ \text{Built-up volume} = \text{Mean area} \times \text{number of buildings} \times \text{mean height} \]

So, in order to determine the buildings’ volumes, it is necessary to know the height of buildings. Thus, we have applied the following method. Since we are using an RGB image and based on the hue range of the image, the green areas have been detected and selected. Then, this RGB image has been transformed into “L-A-B” format and green areas has been made as luminescent as possible to get rid of them in the thresholding procedure. Using Otsu’s method, a multi-threshold image function has been used to divide the image into four categories: highly bright, bright, dark and highly dark where highly dark area is the most probable area to represent the buildings’ shadow. After that, a binary image has been generated and the areas which don’t represent a building’s shadow are removed (Otsu, 1979). At this point, the shadow’s areas are located on the image and the buildings’ heights are determined using the following method: There is a relation between the azimuth of the sun, the dimension of shadow and the height of building, consequently, since the azimuth of sun is known (almanac data) and the dimension of shadow is known (form the already generated image), the building height can be calculated. At the end, the areas and heights of buildings are calculated; the volume can be calculated by multiplying them together.

3.6. Population Data

The population data relative to the chosen villages are determined by visiting each municipality and count the number of unit residence and we got the mean number of family size, so by multiplying the mean number of family size with number of residential unit we get the population in each municipality.

4. Results

4.1. The Volume of Buildings

Using the previously described method, the number of residential buildings (Table 2), the areas of residential buildings and their volumes have been calculated for each municipality (Table 3).
Table 2: Distribution the number of the buildings according to the intervals of surfaces

| Municipality          | Number of buildings between 30-230 m² | Number of buildings between 230-1000 m² | Total number |
|-----------------------|--------------------------------------|----------------------------------------|--------------|
| Aabba                 | 973                                  | 107                                    | 1080         |
| Aain Baalbek          | 810                                  | 193                                    | 1003         |
| Aakkar El-Aatiqa      | 466                                  | 443                                    | 909          |
| Aaytanit              | 21                                   | 33                                     | 54           |
| Amioun                | 307                                  | 349                                    | 656          |
| Babiliyé              | 420                                  | 120                                    | 540          |
| Baskinta              | 586                                  | 238                                    | 824          |
| Beit Chabab           | 600                                  | 213                                    | 813          |
| Beit Meri             | 459                                  | 621                                    | 1080         |
| Bikfaya               | 183                                  | 257                                    | 440          |
| Blida                 | 398                                  | 219                                    | 617          |
| Bqerzla               | 167                                  | 66                                     | 233          |
| Deir Ez-Zahrani        | 538                                  | 404                                    | 942          |
| Dibbabiyé             | 173                                  | 14                                     | 187          |
| Ehden                 | 428                                  | 476                                    | 904          |
| Jbal El-Botm          | 299                                  | 118                                    | 417          |
| Jebaa                 | 34                                   | 12                                     | 46           |
| Joun                  | 352                                  | 137                                    | 489          |
| Khirbet Rouha         | 316                                  | 286                                    | 602          |
| Laboué                | 1193                                 | 245                                    | 1438         |
| Majdel Aanjar         | 658                                  | 608                                    | 1266         |
| Markaba               | 289                                  | 144                                    | 433          |
| Michmich Jbayl        | 190                                  | 74                                     | 264          |
| Nabatieh El-Faouka    | 751                                  | 353                                    | 1104         |
| Qaa Baalbek           | 678                                  | 166                                    | 844          |
| Qalamoun              | 349                                  | 219                                    | 568          |
| Ras El Matn           | 40                                   | 311                                    | 351          |
| Roum                  | 161                                  | 90                                     | 251          |
| Saksakiyé             | 887                                  | 155                                    | 1042         |
| Terbol Zahlé          | 281                                  | 252                                    | 533          |

Table 3: The areas and volumes of buildings for each municipality

| Municipality          | Total number of building | Mean surface (m²) | Buildings surface (m²) | Building volume (m³) | Population |
|-----------------------|--------------------------|-------------------|------------------------|----------------------|------------|
| Aabba                 | 1080                     | 144               | 155520                 | 544320               | 5000       |
| Aain Baalbek          | 1003                     | 171               | 171513                 | 600295.5             | 8000       |
| Aakkar El-Aatiqa      | 909                      | 265               | 240885                 | 963540               | 15500      |
| Aaytanit              | 54                       | 203               | 10962                  | 43848                | 800        |
| Amioun                | 656                      | 280               | 183680                 | 642880               | 7000       |
| Babiliyé              | 540                      | 185               | 99900                  | 349650               | 4000       |
| Baskinta              | 824                      | 204               | 168096                 | 588336               | 7500       |
| Beit Chabab           | 813                      | 178               | 144714                 | 578856               | 16000      |
| Beit Meri             | 1080                     | 288               | 311040                 | 1244160              | 15500      |
| Bikfaya               | 440                      | 298               | 131120                 | 458920               | 7000       |
| Blida                 | 617                      | 240               | 148080                 | 518280               | 4500       |
| Bqerzla               | 233                      | 197               | 45901                  | 137703               | 1500       |
| Deir Ez-Zahrani        | 942                      | 241               | 227022                 | 908088               | 15000      |
| Dibbabiyé             | 187                      | 140               | 26180                  | 78540                | 880        |
| Ehden                 | 904                      | 310               | 280240                 | 1120960              | 16500      |
| Jbal El-Botm          | 417                      | 199               | 82983                  | 248949               | 1350       |
| Jebaa                 | 46                       | 203               | 9338                   | 28014                | 500        |
| Joun                  | 489                      | 198               | 96822                  | 338877               | 4800       |
4.2. Relationship between Buildings and Population

After detection of residential buildings and after we determine the number of building, the area of residential buildings and the volume of buildings, a regression model was created based on census 2013 population data and independent variables for residential building number (Figure 5), area (Figure 6) and volume (Figure 7). The result shows a strong relation with r-squared 0.87.

**Figure 5:** The relationship between the number of buildings and population data

**Figure 6:** The relationship between building areas and population data
The results of this study have been shown a strong relationship between the characteristics of buildings (number, area and volume) and the population data.

To verify the result, a municipality where the numbers of residential and non-residential buildings are known has been chosen. After applying the method used to identify the residential building, the results are summarized in the following confusion matrix (Table 4).

**Table 4:** Confusion matrix of residential building detection accuracy

| Actual situation | Predicted    |
|------------------|--------------|
| Residential building | 901          |
| Non-residential building | 116          |
| Non-residential building | 31           |
| Residential building   | 161          |

\[ \text{Sensitivity} = \frac{TP}{TP + FN} = \frac{901}{901 + 116} = 0.8859 \]

\[ \text{Specificity} = \frac{TN}{TN + FP} = \frac{161}{161 + 31} = 0.8385 \]

The results show that 88.59% of the residential buildings and 83.85% of non-residential buildings were correctly identified.

As a result, the validation of the results and the analysis of error have proved that the margin of error is acceptable. Consequently, the established mathematical relationship can be considered as a valid method to estimate the population data from remote sensing data.

However, this study was limited by the difficulty to identify mixed-use buildings.

### 6. Conclusion

The combination of remotely sensed data and GIS tools was helpful and efficient in building detection and population estimation. In fact, the detection of residential buildings using HSI transformation and “Convex Hull” tool and based on statistical information seemed to be accurate and effortless. Moreover, the estimation of buildings' height is a reliable method for buildings' volume calculation.
taking into consideration the amount of information that might ensure. Furthermore, the established mathematical relationship between buildings’ volume and population allow planners to accurately estimate population using building volumes. At the end, this method might be applied in a variety of situations such as unorganized urban areas and rural areas in order to grasp the necessary information for planners before the establishment of plans or as a monitoring tool for urban sprawl and growth.

References

Council for development and Reconstruction. 2005. National physical master plan of the Lebanese territory. Available from: http://www.cdr.gov.lb/study/sdatl/English/NPMPLT.PDF.

Ghaffarian, S. and Ghaffarian, S. 2014. Automatic building detection based on supervised classification using high resolution google earth images. ISPRS Technical Commission III Symposium, 5 – 7 September 2014, Zurich, Switzerland.

Güdücü, H.V. 2008. Building detection from satellite images using shadow and color information. Available from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.632.5114&rep=rep1&type=pdf.

Guo, T. and Yasuoka, Y. 2002. Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas. Proceedings of 23rd Asian Conference on Remote Sensing.

Inglada, J. 2007. Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features. Journal of Photogrammetry and Remote Sensing, 62(3), pp.236-248.

Jin, X. and Davis, C.H. 2005. Automated building extraction from high-resolution satellite imagery in urban areas using structural, contextual, and spectral information. Journal on Applied Signal Processing, 14, pp.2196-2206.

Koç, D. and Türker, M. 2005. Automatic building detection from high resolution satellite images. International Conference on Recent Advances in Space Technologies. RAST2005, pp.617-622.

Lari, Z. and Ebadi, H. 2007. Automated building from high-resolution satellite imagery using spectral and structural information based on artificial neural networks. Conference on Information Extraction from SAR and Optical Data, with Emphasis on Developing Countries.

Lee, D.S., Shan, J. and Bethel, J.S. 2003. Class-guided building extraction from Ikonos imagery. Photogrammetric Engineering and Remote Sensing, 69(2), pp.143-150.

Rüther, H., Martine, H.M. and Mtalo, E.G. 2002. Application of snakes and dynamic programming optimization technique in modeling of buildings in informal settlement areas. Journal of Photogrammetry and Remote Sensing, 56, pp.269-282.

Tupin, F. and Roux, M. 2003. Detection of building outlines based on the fusion of SAR and optical features. Journal of Photogrammetry and Remote Sensing, 58, pp.71-82.

UN Habitat. 2010. Lebanon – Urban Issues, United Nations Human Settlements. Available from: https://unhabitat.org/lebanon/lebanon-urban-issues/.

Unsalan, C. and Boyer, K.L. 2005. A system to detect houses and residential street networks in multispectral satellite images. Computer Vision and Image Understanding, 98, pp.423-461.