Accuracy Assessment of Large-Scale Topographic Feature Extraction Using High Resolution Raster Image and Artificial Intelligence Method

O Marena¹, M S N Fitri¹, O A Hisam², A Kamil¹, Z M Latif¹,
¹ Department of Survey and Mapping Malaysia, Jalan Sultan Yahya Petra, 50578
Kuala Lumpur, Malaysia
² Faculty of Built Environment and Surveying, University of Technology Malaysia
81310 Skudai Johor, Malaysia

Abstract: Feature extraction is the most important application in spatial data management for updating of topography databases. The importance of features extraction from high resolution raster imagery arises from the fact that it enhances the efficiency of map generation and map production. Presently many studies have been addressed to extract topographic using semi-automated methods using various types of technologies. The objective of the study is to evaluate the planimetry accuracy of topographic features extracted using combination of high resolution and artificial intelligence algorithm. Conventional methodology for features extraction is time consuming and costly. Therefore, automated methods are crucial needed for efficient large area mapping with cost efficient. Maxar 0.3 cm high resolution raster imagery is an important issue in many application domains. The Maxar-Ecopia method is applied in this study to update the large-scale topographic map. Maxar-Ecopia method demonstrates the potential use of high-resolution raster imagery as main source and artificial intelligence algorithm for topographic feature extraction. The planimetric accuracy (RMSEr) of vector extraction is 0.882m. This method is highly accepted to be applied for rapid updating of large-scale topographic map.

1. Introduction

Topographic features are important information for diverse applications. Topographic details need to be mapped because its details are needed for various sectors such as economic development, security, facilities and others. Conventional feature extraction normally is executed manually using digitization and classification process by operator. This method is time consuming and need a lot of manpower. To overcome these limitations, the automatic extraction method provides faster production of vector data. Several studies have been conducted in relation to building extraction from high resolution satellite imagery using object-based image analysis. One of the first studies was carried out by Hofmann in 2001 and followed by many intensive studies on automated feature extraction worldwide such as Maxar (2020), Deepakrishna et.al (2016), Nitin et.al (2016) and Yongyang et.al (2018). Limiting factors for mapping are the information content and the geometric accuracy. General specifications of topographic maps should have a geometric accuracy of 0.2 mm up to 0.3 mm in the published map scale, corresponding to 1.0 m up to 1.5 m for the map scale 1:5000 and 2.0 m up to 3.0 m for the scale 1:10,000.
The limiting factor for topographic mapping is the information content, i.e. the performance of feature extraction from the images. Information content carried out by the image depends also on the characteristics of images, such as effective GSD and used methods for data acquisition - manual and automatic feature extraction.

Therefore, this paper is intended to present the accuracy of the automatic extraction approach on natural and man-made details from orthophoto and high-resolution satellite imagery. In this study, orthophoto and high-resolution satellite images were used to classify different topographic features such as building footprint and transportation categories. The spatial information is divided into relatively homogeneous polygons and extracted according to the detail features information in GIS format.

2. Feature Extraction Approach

Feature extraction is an essential process for image data dimensionality reduction and classification. However, feature extraction is very difficult and often requires human intervention. Artificial intelligence algorithms can achieve automatic feature extraction and image classification but the majority of existing methods extract low-level features from raw images without any image-related operations. Furthermore, the work on the combination of image-related operators/descriptors for feature extraction and image classification is limited. There are various methods of automatic feature extraction that have been developed for example pixel-based algorithms, object-based algorithms, landbase, tensor flow, neural network, shadow and cognition. This paper is focusing on Maxar-Ecopia feature extraction methodology. Maxar-Ecopia is an artificial intelligence and machine learning approach that creates high definition (HD) vector maps. This approach is innovatively developed to speed up the production of vector maps with human accuracy in less time and cost effective than traditional methods.

2.1 Maxar-Ecopia Technology

Ecopia and Maxar have engineered the first high-precision, semi-automated building footprints product by integrating Maxar’s industry-leading high resolution satellite imagery, Ecopia’s advanced artificial intelligence and the cloud compute power of GBDX (Maxar, 2020). Ecopia has defined a number of quality assurance thresholds which represent the quality of Ecopia Surface Map. Some of these quality assurance thresholds are not feasible to verify against the entire dataset, for these areas a statistical sampling is applied or in order to ensure the quality of the data for aspects that cannot be globally verified, the following sampling methodology will be applied (Table 1 and Figure 1):

| Measure Name       | Threshold | Sample/All Polygons |
|--------------------|-----------|---------------------|
| False negatives    | <=5%      | Sample              |
| False positives    | <=5%      | Sample              |
| Valid interpretation| >95%      | Sample              |
| Valid geometry     | 100%      | All polygons        |
| Minimum area       | 100%      | All polygons        |
2.2 Source of Data

The primary datasets for this study comprise of two (2) high resolution raster images. Maxar 0.3 m and 0.15 m spatial resolution for satellite imagery and orthophoto, respectively. Maxar 0.3 m spatial resolution provides high planimetric accuracy (better than 5 m CE90) closely aligned to known coordinates on the surface of the earth (Maxar, 2020(b)). Figure 3 shows the sample of 0.3 m spatial resolution of Maxar high resolution satellite imagery. The Department of Survey and Mapping Malaysia (DSMM) orthophoto is an aerial photograph that has been geometrically corrected or ortho-rectified with map scale 1:5,000 in GDM 2000 RSO coordinate system (Figure 2).

2.3 National Topographic Database

The National Topographic Database (NTDB) shows planimetric details and altimetric details. NDTB databases portray detailed ground relief, drainage, forest cover, administrative zones, built environment and other man-made features, NTDB consists of map series (1:5,000 to small scale 1:50,000) showing 12 data categories as shown in Figure 3 (MaCGDI, 2010).

Figure 1. Building Footprints and Road Features Output From Maxar-Ecopia Landbase (Langkawi Island, Malaysia)

Figure 2. (Left) Maxar 0.3 m Spatial Resolution Satellite Imagery (Langkawi and Penang) and (Right) 0.15m Spatial Resolution of Orthophoto

Figure 3. Malaysian Standard Geographic Information/Geomatics Feature Codes and Attributes (MS1759)
3. Implementation Methodology

The methodology will discuss in detail on the method and strategy used to collect and process the datasets according to study implementation. Figure 4 presents the methodology of planimetric accuracy assessment of raster image and vector data generated from automated feature extraction process. Ground control points (GCP) and check points (CP) were established for image rectification process on acquired original rectified raster image. This step is to improve the accuracy of the image. Root Mean Square Error (RMSE) values obtained for GCPs and CPs were used to determine the planimetric accuracy of these images. Root mean square (RMS) error of image rectification is verified before it can be used for vector spatial adjustment. Check point (CP) is utilize to verify the planimetric accuracy of orthorectified raster and vector for study area. Location of CPs are identified randomly with respect to orthorectified image. The selected locations of CPs include man made structure, road junction and road marker. Comparison of coordinate between CP and adjusted vector (based on orthorectified image) will indicate the planimetric accuracy whether the adjusted vector is complying with Malaysian Standard ISO 9001:2005 (Figure 4).

![Figure 4. Positional Accuracy Evaluation Methodology](image)

4. Planimetric Accuracy Evaluation

Requirements for the accuracy of geometric correction of raster imagery vary depending on the application. In the case of satellite and aerial based mapping, the requirements can be obtained from base mapping accuracy standards. The coordinates of generated points are then compared with points determined by aerial photogrammetry or satellite imagery. From the differences, accuracy measures are derived. Tests regarding the type of error distribution are carried out in order to apply either standard accuracy measures or robust accuracy measures. The results of a practical test are used to assess the planimetric accuracy by means of various accuracy measures. Maxar 0.3m resolution satellite image (Maxar, 200b) specification tabulates the accuracy of less than 4.2 m with (CE90) circular error at the 90th percentile (CE90). A total of 42 GCPs and 11 CPs have been established in the study area (Figure 5) covering states of Penang and Kedah.

For the purpose of preliminary planimetric accuracy checking of satellite imagery, 15 GCP stations have been selected randomly in the study area. The minimum and maximum coordinate differences in easting and northing components is ranging from 0.146 m to 2.038m while RMS for northing (RMSEy) and easting (RMSEx) components is 0.544 m and 1.082 m, respectively. Table 2, demonstrates that planimetric accuracy of satellite imagery for study area is better and following Maxar (200b) specification. The establishment of control points (GCP) allows the rectification process to be implemented locally. This allows planimetric accuracy of 0.3m resolution satellite imagery for the study area to be improved compare to the original specification. In order to improve the accuracy of satellite imagery coordinate than second rectification process is executed using 42 GCP control points (Figure 5).
A total of 11 check point (CP) stations were used for planimetry accuracy analysis. Table 3 tabulates the result of second rectification. The minimum and maximum coordinate differences in easting and northing components is ranging from 0.033 m to 1.156 m while RMS for northing (RMSEy) and easting (RMSEx) components is 0.499 m and 0.766 m, respectively. ASPRS (2013) accuracy standards for digital geospatial data stated that the tolerance for 1st class digital planimetric data horizontal accuracy/quality (RMSEr) is 0.884 m for 1:5,000 map scale. From this study the total RMSEr is 0.882m. RMSEr equals the horizontal radial RMSE, i.e., √(RMSEx² + RMSEy²). All RMSE values and other accuracy parameters are in the same units as the pixel size. For example, if the pixel size is in m, then RMSEx, RMSEy, RMSEr, horizontal accuracy at the 95% confidence level, and seamline mismatch are also in meters.

**Table 2. Coordinate Differences Between Rectified Image and GCP**

| Point Name | GCP Coordinate | Rectified Coordinate | Differences |
|------------|----------------|----------------------|-------------|
|            | N (m) | E (m) | N (m) | E (m) | Diff N (m) | Diff E (m) |
| GP04       | 595198.733 | 256525.598 | 595198.539 | 256525.314 | 0.194 | 0.284 |
| GP05       | 598987.545 | 255005.963 | 598987.598 | 255004.904 | -0.053 | 1.059 |
| GP014      | 600427.367 | 267813.626 | 600427.197 | 267812.897 | 0.170 | 0.729 |
| GP015      | 602927.939 | 265809.506 | 602927.635 | 265,809.65 | 0.304 | -0.143 |
| GP025      | 614808.789 | 275316.161 | 614808.714 | 275316.286 | 0.075 | -0.125 |
| GP026      | 614563.425 | 279879.84 | 614563.422 | 279880.842 | 0.003 | -1.002 |
| GP027      | 610019.443 | 284588.75 | 610020.130 | 284588.870 | -0.687 | -0.120 |
| GP033      | 621705.91 | 277342.556 | 621705.891 | 277343.900 | 0.019 | -1.344 |
| GP034      | 624904.611 | 279966.11 | 624904.435 | 279967.461 | 0.176 | -1.351 |
| GP036      | 624612.746 | 283698.565 | 624611.310 | 283699.013 | 1.436 | -0.448 |
| GP037      | 629037.316 | 275764.988 | 629036.385 | 275766.054 | 0.931 | -1.066 |
| GP038      | 629757.561 | 279739.006 | 629757.113 | 279740.084 | 0.448 | -1.078 |
| GP041      | 604972.96 | 275359.06 | 604972.928 | 275357.499 | 0.032 | 1.561 |
| GP042      | 598052.27 | 267894.51 | 598051.518 | 267892.616 | 0.752 | 1.894 |
| GP043      | 626539.386 | 277799.718 | 626539.129 | 277801.341 | 0.257 | -1.623 |

Min -1.623 0.146
Max 1.894 2.038
RMSE 0.544 1.082
4.1 Geometry Accuracy of Extracted Vector Data

Feature extraction is an approach to extracts important regions or objects of interest algorithmically from large data sets. In order to quantify the accuracy, reference data are required then calculate a radial error or metric, which measures the similarity between the result and the reference data. Reference data can be acquired through manual digitation of visually extractable linear features or perform overlapping data on images. For this study, the accuracy assessment is conducted by comparing the extracted vector with high resolution raster image. Figure 6(a) show the overlay analysis between extracted vector (building footprint) and high-resolution raster image. Figure 6 (b) and Figure 6 (c) visualizes the accuracy of extraction process using artificial intelligence method.

Figure 6 (a). Overlay Analysis Between Extracted Vector (AI) and Raster Image At Maximum Zoom Level (1:150-1:400).

Figure 6 (b). Overlay Analysis Between Extracted Vector (AI) and Raster Image For Building Footprint.

Table 3. Coordinate Differences Between Second Local Rectified Image and CP

| Point Name | CP Coordinate | Local Rectified Image Coordinate | Differences |
|------------|---------------|----------------------------------|-------------|
|            |  N (m) | E (m)  |  N (m) | E (m)  |
|------------|--------|--------|--------|--------|
| CP01       | 592048.684 | 257184.896 | 592048.863 | 257184.295 | 0.179 | 0.601 |
| CP02       | 598451.556 | 257279.468 | 598451.617 | 257278.482 | 0.061 | 0.986 |
| CP03       | 603196.227 | 257977.022 | 603195.439 | 257977.553 | 0.788 | 0.531 |
| CP04       | 596567.588 | 265854.023 | 596567.828 | 265854.111 | 0.280 | 0.081 |
| CP05       | 601217.318 | 266371.395 | 601218.123 | 266370.345 | 0.805 | 1.050 |
| CP06       | 602120.977 | 276151.574 | 602121.010 | 276151.806 | 0.033 | 0.232 |
| CP07       | 607358.556 | 277072.097 | 607359.530 | 277071.281 | 0.974 | 0.816 |
| CP08       | 613396.566 | 282619.246 | 613396.856 | 282618.140 | 0.290 | 1.106 |
| CP09       | 623337.166 | 278058.722 | 623337.661 | 278059.878 | 0.495 | 1.156 |
| CP10       | 622556.159 | 282556.296 | 622556.362 | 282555.792 | 0.203 | 0.504 |
| CP11       | 626978.631 | 279779.362 | 626978.381 | 279779.798 | 0.229 | 0.436 |

Min: -0.974, 1.156
Max: 0.788, 1.106
RMSE: 0.499, 0.766
Based on the results in Figure 6 (b) and Figure 6 (c), the accuracy of AI vector extraction is confirmed to ASPRS (2013) which specified that horizontal accuracy at the 95% confidence level for 30 cm orthophoto or satellite imagery is 0.734 m.

5. Conclusion

This exploratory study aimed to explore an alternative method for rapid large scale topographic feature extraction and topographic database generation and production. The applicability of combination of high-resolution raster imagery and AI on automated and semi-automated feature extraction provides an exemplary output product. The most common accuracy assessment approaches were described. In conclusion, this paper serves as a basis for the subsequent implementation of suitable methods in automated feature extraction using high resolution raster imagery and AI method workflow for large scale topographic map scale 1:5,000. Due to a lack of robustness of automatic feature extraction, semi-automatic approaches that combine the interpretation skills of a human operator is needed. This approaches that include editing capabilities seem indispensable for topographic database. The approach should correctly and completely extract all relevant topographic features, be simple in parameterization with a high degree of automation and a minimal need of interaction, self-assessment to increase reliability and be robust against varying quality of input data. Furthermore, the combination of Maxar 0.3 m high-resolution raster imagery and Ecopia (AI) is revitalizing the instrument of rapid and updated topographic database generation and map publication.

References

[1] American Society for Photogrammetry and Remote Sensing, 2013. Accuracy Standards for Digital Geospatial Data. United States of America.
[2] Deepakrishna, Ediriweera and Jagath Gunatilake, 2016. Automatic Feature Extraction From Satellite Images Using Lvq Neural Network. 7th Asian Conference on Remote Sensing, Colombo.
[3] Ecopia, 2020. Ecopia Tech: Digitizing The World Using AI. Ecopia Tech, Toronto.
[4] Hofmann, P., 2001. Detecting Urban Features From IKONOS Data Using An Object-Oriented Approach. First Annual Conference of the Remote Sensing & Photogrammetry Society (RSPS2001) Proceedings. Munich, Germany.
[5] Department of Survey and Mapping Malaysia, 2005. Malaysian Standard ISO 9001:2005. Ministry of Natural Resources and Environment, SIRIM. Shah Alam.
[6] Malaysian Centre for Geospatial Data Infrastructure, 2020. MyGDI: The Enabling Platform Towards Spatially Enabled Government In Malaysia. Kementerian Sumber Asli dan Alam Sekitar, Putrajaya.
[7] Maxar, 2020. Ecopia Building Footprints Powered By Maxar. Maxar Corporation. United States of America.
[8] Maxar, 2020b. Earth Intelligence. Maxar Corporation. United States of America.

[9] Crommelinck, S, Bennett, R, Gerke, M.N, Francesco, Y.Y, Michael. G, Vosselman, 2016. Review of Automatic Feature Extraction from High-Resolution Optical Sensor Data for UAV-Based Cadastral Mapping. Remote Sensing, 8, 689. MDPI, Switzerland.

[10] National Research Council, 2003. Weaving A National Map. National Academies Press. Washington.

[11] Nitin L. Gavankar & Sanjay Kumar Ghosh, 2018. Automatic Building Footprint Extraction From High Resolution Satellite Image Using Mathematical Morphology. Geocarto International. Taylor & Francis, UK.

[12] Yongyang Xu, Liang Wu, Zhong Xie and Zhanlong Chen (2018) Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters. Remote Sens, MDPI. Switzerland