An Approach for Evaluating the Success of Information Extraction Procedures on Area Features

Srivastava, V.
Geoinformatics Department, Indian Institute of Remote Sensing, Indian Space Research Organisation, Department of Space, Govt. of India, Dehradun, India, E-mail: vanditasr@gmail.com, vandita@iirs.gov.in

Abstract
The present study proposes a conceptually simple yet effective approach for evaluating the success of information extraction procedures on area features, using five parameters: spatial error of omission, spatial error of commission, correctness, completeness and quality. To demonstrate this approach, three region based segmentation algorithms were applied to semiautomatically extract the outlines and areas of buildings from high-resolution satellite imagery. Qualitative and quantitative comparisons of the segmentation results demonstrate that the proposed parameters are able to evaluate the success of information extraction processes on area features. The five proposed parameters, supplemented with a shape index can be useful for evaluating result of information extraction processes on area features in all fields dealing with extraction of natural and manmade features.

1. Introduction

With the launch of high-resolution (H-R) satellite sensors offering high spatial and radiometric resolution, it has become easy to visually assess fine-resolution natural and man-made features for applications such as topographic mapping (Holland et al., 2006). However, manual extraction of features from the H-R imagery over large areas is tedious and time consuming. Various researchers have developed automatic information extraction methods applied to H-R data (Ahmadi et al., 2010, Giada et al., 2003, Bacher and Mayer, 2005, Brekke and Solberg, 2005, Inglada, 2007, Xu et al., 2002 and Collins et al., 1995). Blaschke (2010) provides an extensive review of object-oriented segmentation methods for extracting area features. However, selecting the optimum method in a given situation is problematic. The analyst faced with this task, as well as the ultimate client of information produced by automatic extraction, requires a quantitative evaluation of the success of each approach. Numerous evaluation methods have been developed for evaluation of information extraction processes such as segmentation (Polak et al., 2009, Zhang et al., 2008, Udupa et al., 2006 and Zhang, 1997). These methods are comprehensive but complex. Therefore, there is a need for a simple yet valid methodology for evaluating, both spatially and statistically, the success of information extraction processes which result in area features. Neubert and Meinel (2003) used average difference in area, perimeter, shape index and visual quality as the parameters for evaluating result of segmentation processes which produce area features. The encouraging results obtained by these authors motivate the search for improvements.

1.1 Objectives

This study has two objectives: (1) to develop evaluation criteria for quantitative evaluation of the success of extraction of area features; (2) to use these to compare several popular segmentation algorithms.

1.2 Data used

A panchromatic IKONOS image (1m ground resolution, 11 bit radiometric resolution) of Dehradun City, Uttarakhand State, India (30°19′N and 78°20′E), (Figure 2) acquired on 19th April, 2001 was used as a test image, because of the ease of ground truthing. A single representative test site (400 by 280 pixels) was selected for analysis, having all the desired features: free-standing buildings, buildings intermixed with vegetation, and paved areas, thus providing a challenge to the extraction process).

2. Methodology

2.1 Evaluation Methodology

Accuracy is the degree of conformity with a true reference. Wiedemann et al., (1998) have described algorithms to check accuracy. Accuracy exhibits different parameters like *completeness, correctness, quality*. The concept of using three parameters for linear features proposed by Wiedmann et al., (1998) was extended to derive parameters for evaluation of extraction on area features. The following five parameters are proposed for spatially evaluating the result of any extraction process on area features (Figure 1):

Spatial error of omission: in case of an area feature, describes which and how much area was omitted from extraction, and can be defined quantitatively as the ratio of the unmatched reference feature area to the total reference area.

Spatial error of omission (S.E.O.): area of unmatched reference/ area of reference; S.E.O. $\in \{0;1\}$

Spatial error of commission: in case of an area feature describes which and how much area was wrongly committed in the extraction, and can be defined quantitatively as the ratio of the unmatched extracted area to the total extracted area.

Spatial error of Commission (S.E.C.): area of the unmatched extraction/ total extracted area; S.E.C. $\in \{0;1\}$

Completeness: of an area feature describes where and how much complete an area layer is and can be defined quantitatively as the ratio of the correctly extracted area to the total reference area.

Completeness: area of matched extraction/ total reference area; Completeness $\in \{0;1\}$

Correctness: of an area feature describes where and how much the area features are correctly extracted and can be defined quantitatively as the ratio of the correctly extracted area to the total area extracted.

Correctness: area of matched extraction/ total area extracted; Correctness $\in$ {0;1}

Quality: of an area feature combines *completeness* and *correctness* to give a measure of final *quality* of the result and can be defined as the ratio of the correctly extracted area to the total area, i.e. sum of area extracted and area of reference that was not extracted.

Quality: area of matched extraction/(area of extraction + area of unmatched reference); Quality $\in$ {0;1}

2.2 Segmentation Algorithms used

Results of building extraction using three region based segmentation algorithms were evaluated using the parameters proposed in Section 2.1.

The three segmentation algorithms have been implemented using development platform of three popular software packages commonly used for satellite and aerial data processing of earth resources (viz., ENVI/RSI, eCognition and ERDAS/Imagine). These algorithms are described by their respective developers but are not explicitly named, so we refer to them as Region Segmentation Based on DN Range (RSBDNR), Bottom Up Region Merging Approach (BURMA) and Distance Based Segmentation Approach (DBSA) respectively. The present section briefly describes the segmentation algorithms used by each package.

1. Region Segmentation Based on DN Range (RSBDNR): The first segmentation algorithm is a simple region-based approach, developed by Research Systems Inc., USA, which segments the image into areas of connected pixels based on the pixel DN value, or a range of DN values (ENVI user Guide, Sept 2003, RSI, USA). This means that only pixels that fall within the entered DN range will be considered in making the segmentation image. All other pixels will have an output value of 0. Either four or eight adjacent pixels are considered for the connectivity to form a region. A region has to follow the criteria of minimum number of pixels specified. Each connected region, or segment, is given a unique DN value in the output image. If only one value is entered, the data minimum or maximum is used as the other end of the threshold.
2. Bottom Up Region Merging Approach (BURMA): The second approach is a multi-resolution segmentation approach developed by Definiens Imaging GmbH, Munich Germany, which uses a bottom-up region merging technique to extract homogeneous image object primitives at a chosen resolution (eCognition 4, User Guide, Definiens Imaging GmbH). Homogeneous image objects are achieved by minimizing weighted heterogeneity using tone and shape as the parameters to calculate heterogeneity. The input parameters to the algorithm are *scale parameter*, layer *weights* and the mixing of the heterogeneity criterion based on *colour* (tone) and *shape*. Adjusting the scale parameter indirectly influences the average object size. Either plane or diagonal definitions of neighbourhood can be used.
3. Distance Based Segmentation Approach (DBSA): This is a simple segmentation approach developed by the USDA Forest Service, Remote Sensing Applications Center, Salt Lake City, Utah (B. Ruefenacht, personal communication, April 2010).

The algorithm is like ISODATA along with the spatial collocation component, which considers Spectral Threshold Euclidean Distance for forming of segments. The main input parameters to the segmentation module are Spectral *Threshold Distance* for limiting the growth of the region and the *Minimum Region Size*, which defines the size for the minimum region. The program selects a pixel and computes the Euclidean distance for an adjacent pixel. If the Euclidean distance is less than the spectral threshold, it is included in the segment. It does this for all 8 adjacent pixels. If a pixel was included in a segment, this pixel is used as the center pixel and all 8 pixels surrounding this pixel are tested to see if they fall within the spectral threshold limit. A larger spectral limit means a larger segment. The minimum region size just eliminates pixels smaller than the size selected. The size is in pixels. The segments to be eliminated are actually merged with the neighboring segment that is the most spectrally similar to the segment under consideration.

2.3 Establishing Optimum Segmentation Parameters for Extracting Buildings

As the three segmentation algorithms are based on parameter tuning, therefore, to compare the result of extraction process, it is important that best (optimum) parameter combinations are used in each algorithm to extract the target features, which enable the best and comparatively, similar extraction results in each case. In order to achieve this set of optimum parameters, several parameter combinations were used on several test images in each of the three algorithms for extraction of building. Table 1 presents the optimum (best) parameters in the three modules for extraction of buildings. During the analysis of segmentation result, a comparison was established amongst parameters of segmentation in the three modules. It was found that *Min Population* parameter of RSBDNR was functioning similar to the *scale parameter* in BURMA and *Min Region size* and *Spectral threshold Distance* in DBSA, all three deciding the size of object in turn. The values of these parameters were kept constant to have uniformity in result. In RSBDNR, the selection of range was done based on the analysis of DN values for building objects. The selection of range was tricky as expanding the range was including unwanted regions, and compression of range, on the other hand, was ignoring some of the building pixels, due to the similarity of DN values of building objects with other features having same tone like that of roads, parks etc. in the PAN image. Building objects with very dark roof tops also caused difficulty in selection of parameters. Minimum size of the building was one important deciding factor in selection of *minimum population*. In case of BURMA, apart from the *scale parameter*, which indirectly decided about the size of the object, which was kept constant as discussed above, the values of parameters *shape* and *compactness* were fixed in combination based upon trial and observation. A balance was achieved first in weights to *tone* and *shape* with view to extract building objects, and thereafter a trade-off between the *compactness* and *smoothness* was decided. DBSA did not give much opportunity to differentiate building objects from other objects except that the range DN values of building objects was analyzed in the test images based upon which an optimum *spectral threshold distance* was decided for extraction of building objects. While RSBDNR was flexible in its segmentation module as it gave opportunity to define the range, BURMA had an additional advantage of utilizing the shape parameter of the objects for performing segmentation. However RSBDNR did not include vectorisation function for obtaining vector objects from pixel regions from the segmented result immediately after segmentation. To avoid non-uniformity in vectorisation results, no smoothening was used during vectorisation, as the BURMA and DBSA had different smoothening criteria.

Table 1: Optimum parameters for extraction of buildings

S. No.	Parameter	RSBDNR	BURMA	DBSA
1.	Size of Objects	50 (Min Population)	50 (Scale Parameter)	50,50 (Min Region Size, Spectral threshold Distance)
2.	Min/Max.	550, 900	-	-
3.	Shape & Compactness	-	0.7, 0.5	-
4.	Block Size	-	-	100

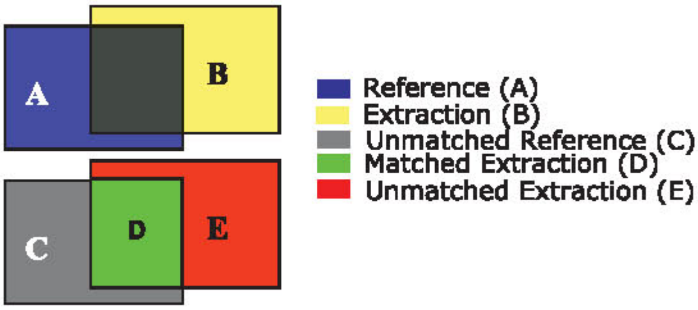


Figure 1: Matching scheme for extraction result w.r.t. reference

Figure 2: Test Image

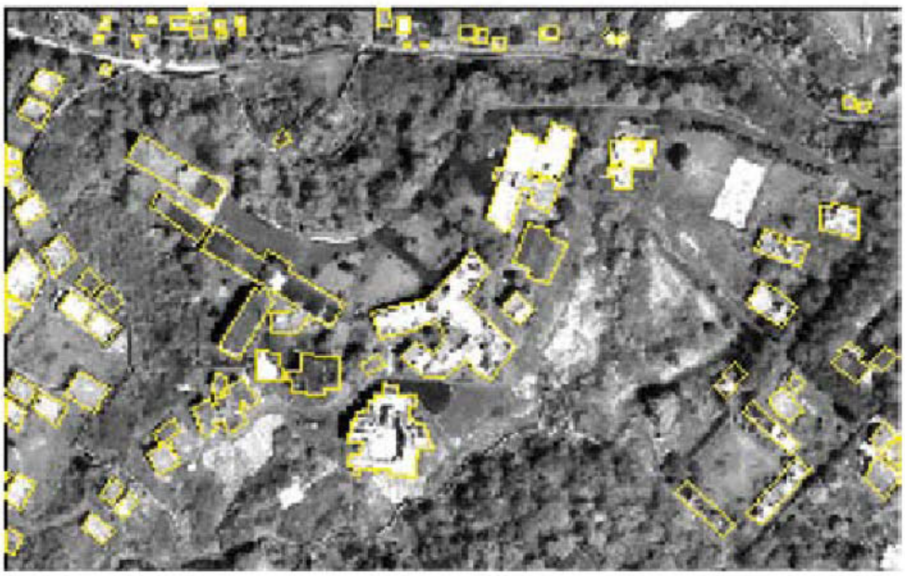


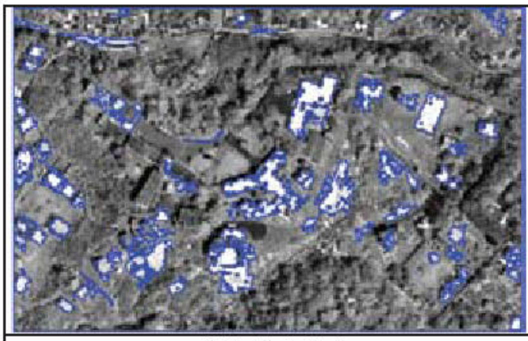
Figure 3: Reference

2.4 Evaluation of Segmentation Result

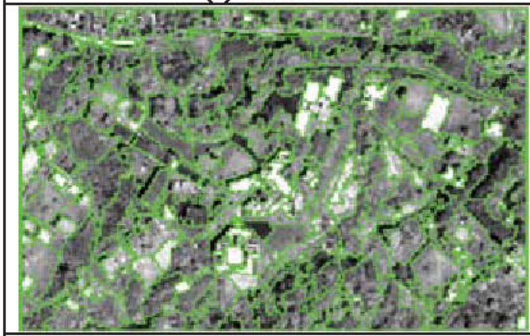
No extraction is complete until its accuracy has been assessed, however misuse of statistics may mislead the accuracy assessment. Therefore it is essential to have a simple and reliable method for accuracy assessment. The accuracy of the segmentation result was checked both qualitatively and quantitatively using the methodology proposed in section 2.1. To perform the evaluation of the segmented results, they were first vectorised and brought to one common platform where first the results were analyzed visually and qualitatively and then a detailed quantitative analysis was performed to judge the quality of segmentation based on proposed method of evaluation, followed by a general comparison of the three algorithms.

2.4.1 Qualitative analysis

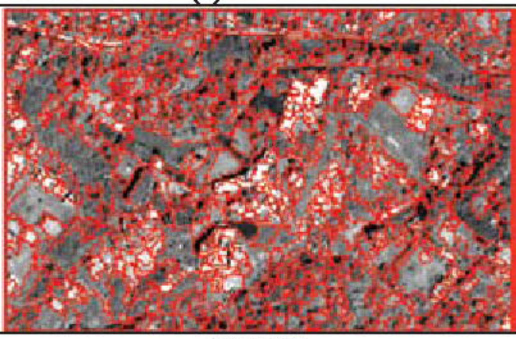
The Qualitative analysis included visual survey with respect to the reference (Figure 3) of the original segmentation result (Figure 4) overlaid over IKONOS (PAN) image, as well as of the segmented result corresponding to buildings extracted with respect to reference layer (Figure 5), overlaid over IKONOS (PAN) image. All the results were compared based on criterion of distinct delineation of buildings from other similar features like roads, footpaths, open ground and vegetation; object shape and size with respect to the building objects in the reference, inclusion of non building objects and exclusion of building objects and also mixing/segregation of objects with respect to corresponding reference building objects.



(a) RSBDNR

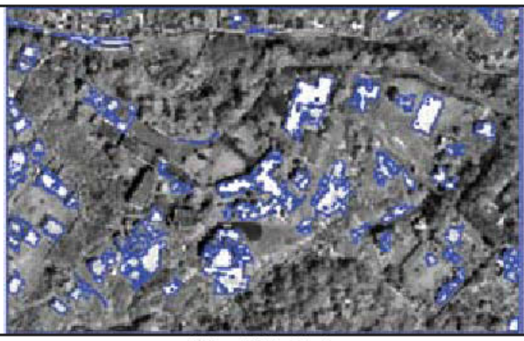


(b) BURMA

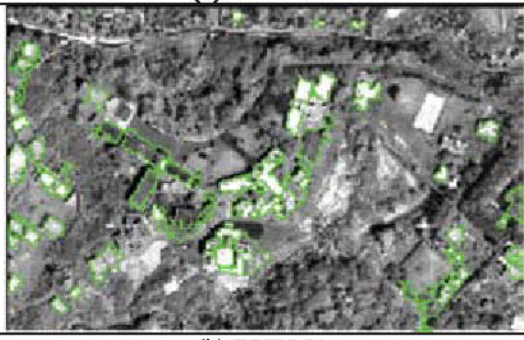


(c) DBSA

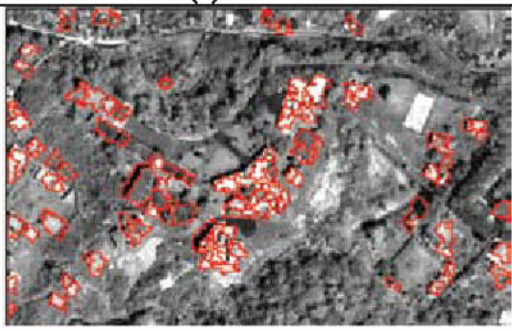
Figure 4: Original Segmentation Result



(a) RSBDNR



(b) BURMA



(c) DBSA

Figure 5: Segmentation Result (Building Objects Only)

2.4.2 Quantitative analysis

Methodology evolved for exhaustive quantitative comparison of segmented result in section 2.1 was used for detailed quantitative comparison of extracted building objects (Figure 5, segmented buildings, overlaid over IKONOS (PAN) image) Subsets were created of the results from there algorithms having only building objects and 14 samples were selected in all the three segmentation results having a correspondence with the reference layer (Figure 6, segmentation result buildings-14 samples, overlaid over IKONOS (PAN) image), their statistics of evaluation was computed based on the proposed evaluation method (Table 2, Figure 5). Figure 7 shows an overall quantitative comparison of the total statistics of selected building objects in the three cases.

Table 2: Quantitative Comparison of buildings extracted from three segmentation algorithms using proposed methodology

Parameter	Reference	RSBDNR	BURMA	DBSA
Completeness **(min, max, average)**	-	0.388, 0.892, 0.708	0.279, 0.705, 0.890	0.468, 0.925, 0.759
Correctness **(min, max, average)**	-	0.590, 0.998, 0.841	0.329, 0.989,0.862	0.597, 0.942, 0.776
Quality **(min, max, average)**	-	0.310, 0.855, 0.624	0.273, 0.876, 0.630	0.389, 0.760. 0.621
Shape Index **(min, max, average)**	105.93, 3107.48, 781.77	142.07, 8406.00, 1990.85	131.58, 4083.01, 975.07	118.44, 5194.38, 1181.48

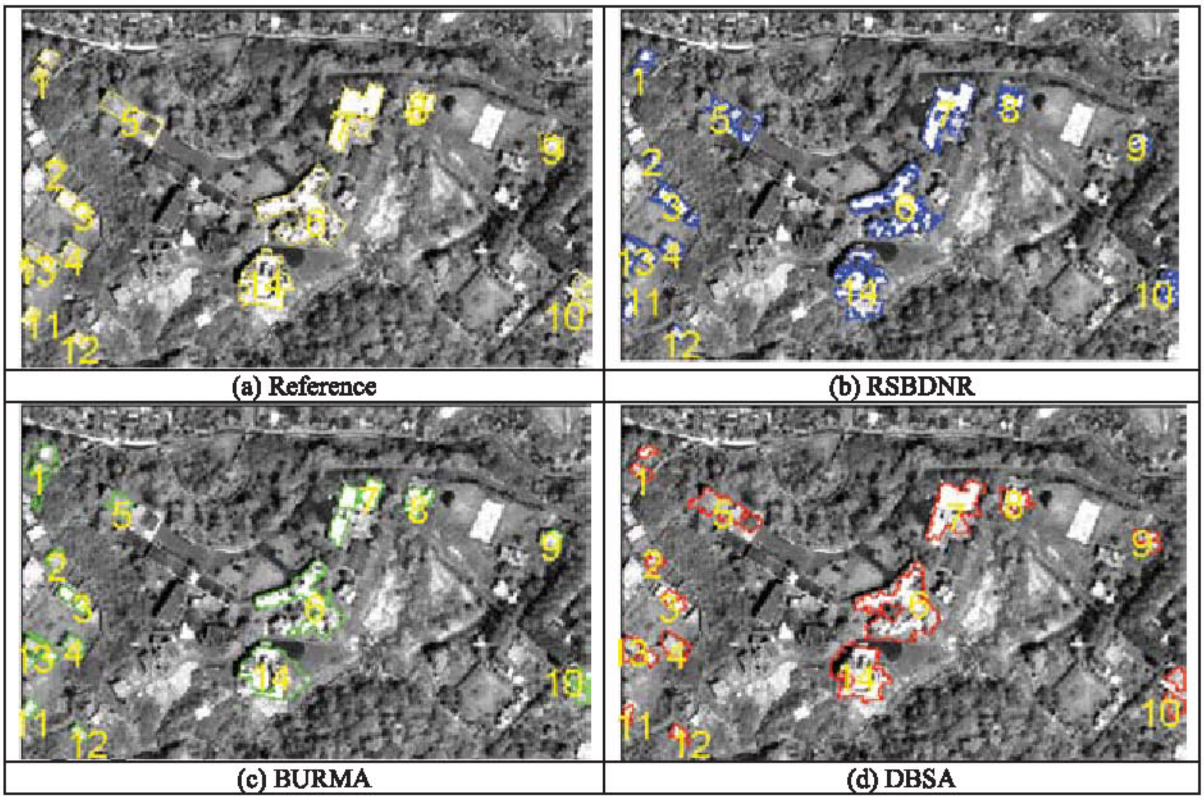


Figure 6: Segmentation Result for selected 14 samples of buildings

2.4.3 Spatial evaluation

In addition to qualitative and quantitative analysis, a spatial evaluation was also carried out. The two parameters namely *Spatial error of omission* and *Spatial error of commission* were used along with matched extraction to prepare three maps, namely *Omission Error Map*, *Commission Error Map* and Quality Map. *Omission Error Map* shows areas of buildings out of total reference area that were omitted from extraction, *Commission Error Map* shows areas that were not part of buildings but were committed to be buildings, and *Overall Quality Map* showed the spatial location of the areas that were extracted correctly (matched extraction) in addition to the areas omitted from extraction (unmatched reference) and the areas committed to be buildings by mistake (unmatched extraction) (Figure 8).

3. Results and Discussion

Several quality parameters were developed to evaluate the result of information extraction process on area features. To demonstrate the usefulness of the proposed evaluation approach, an exhaustive qualitative and quantitative evaluation of results obtained from three segmentation algorithms (described in section 2.2) was done. The IKONOS panchromatic image has been used for the evaluation purpose. A computationally inexpensive approach has been proposed for extracting buildings using image segmentation techniques. The optimum parameters for segmenting buildings have been established. The results of three segmentation algorithms are assessed using the proposed evaluation methodology and the results are discussed in subsequent subsections, finally concluding with a general comparison of the three segmentation modules.

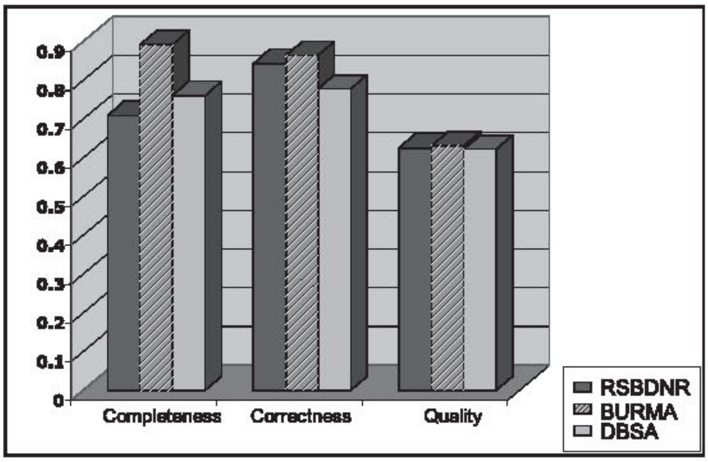


Figure 7: Quantitative Comparison of (average performance of) three segmentation modules

Table 3: General Comparison of three segmentation algorithms

Segmentation Algorithm	RSBDNR	BURMA	DBSA
Algorithm	Region Segmentation Based on DN Range	Bottom Up Region Merging Approach	Distance Based Segmentation Approach
Basis of Segmentation	Tone	Tone, shape	Tone
Output Statistics of Segmented Object	DN Value	Exhaustive statistics (207 features) on Object related features, class related features and global features	Limited (Grid code, ID, area and perimeter)
Parameters	3 (Min, Max Min Population)	3 (Scale Parameter, Shape, Compactness)	3 (Block Size, Spectral Threshold Distance, Min. Region Size)
Max Image size handled by segmentation (pixels)	-	10000 by 10000	2000 by 2000
Retaining of Spatial Referencing Information	Yes	No	Yes
Vectorization of Segmentation Result	No	Yes	Yes
Flexibility during vectorization	Not Applicable	Allows for smoothening	Allows for smoothening

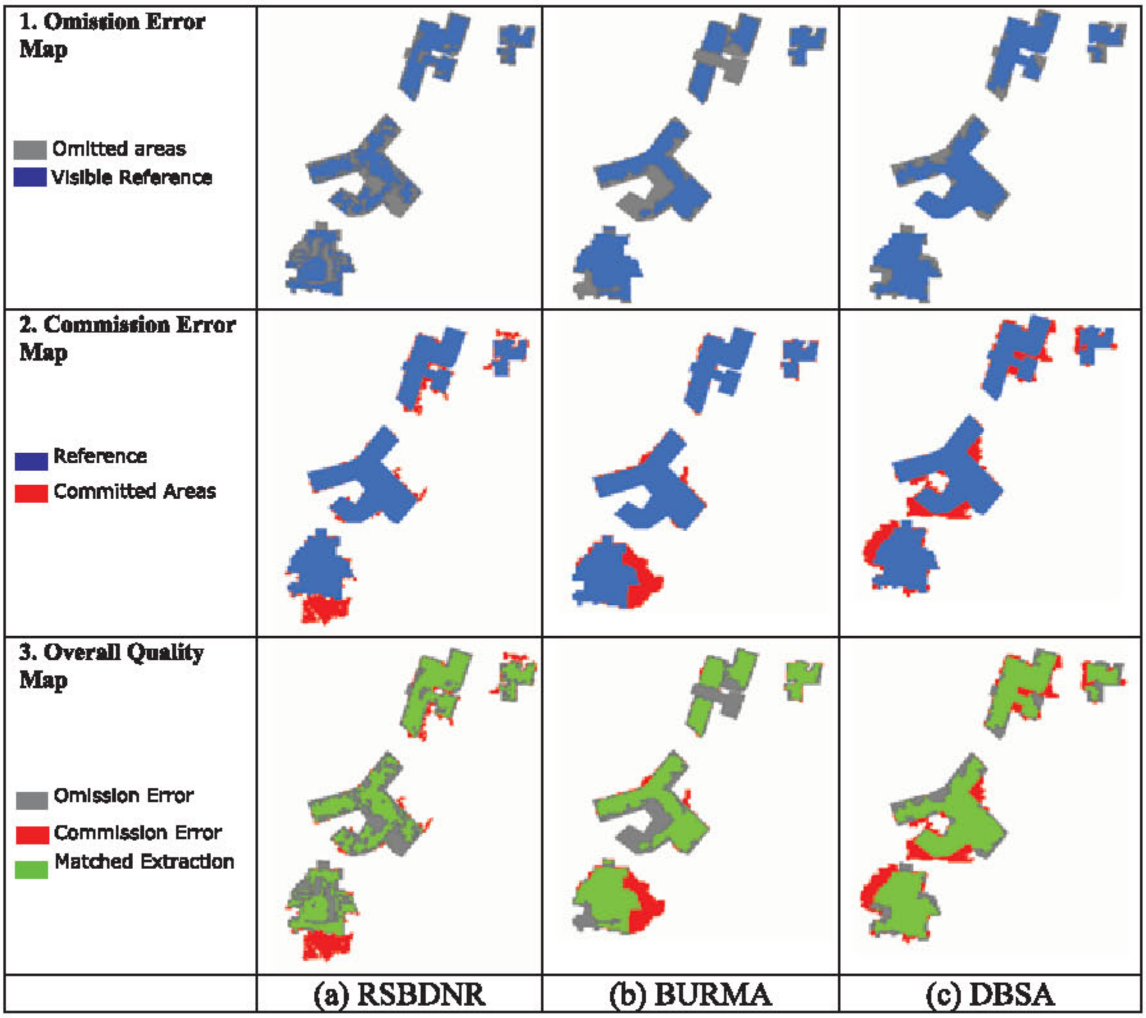


Figure 8: Spatial Evaluation Maps

3.1 Qualitative Analysis

RSBDNR was flexible enough to incorporate the range of pixel values for building objects and most of the building objects could be easily separated from other objects at the segmentation stage itself. Most of the building objects were extracted, except for the cases where the building roofs were dark (due to the shielding by tar sheet to protect from leakage) and was having intensities very less as compared to other building objects (Figure 4a). Some non-building objects such as roads & open grounds were also extracted as buildings due to their mix with the intensities of building objects. In case of BURMA and DBSA (Figure 4b, 4c), it was not possible to differentiate building objects from other objects due to the limitation of choice in segmentation parameters. In case of BURMA, because of the selection of segmentation parameters suitable to building objects, the border of the building objects was smoother than that of RSBDNR and DBSA, and at the places where RSBDNR and DBSA further divided the building object due to the within-object-variability of intensity, BURMA gave better results in terms of size of the object still maintaining shape of its outer boundary. DBSA had difference in its object size, extracting bigger objects where the spatial frequency was low, and smaller objects in case of high spatial frequency, whereas RSBDNR objects were of comparatively smaller size than the reference objects, and had, on an average, the object sizes matching with the reference buildings, except for building objects with highly frayed boundaries.

3.2 Quantitative Analysis

The results of quantitative analysis over 14 selected samples (Figure 6) are summarized in Table 2. The sample objects were selected considering maximum heterogeneity, such as that in size, shape, tone, location. The chart in figure 7 shows that the highest values of all the three variables were observed in case of BURMA. Further, buildings extracted using BURMA was more complete and had a high value of *quality* as compared to other two. The second best performance was of RSBDNR considering the *Correctness* and *Quality*, however the *Completeness* was lowest in RSBDNR. In case of *shape index*, the quantitative analysis shows that on an average BURMA performed better than the other two. Even though the min and max value of *completeness* and *quality* parameters of the three modules depicted that DBSA and RSBDNR performed better, in many cases, however, on an average, for most of the objects, BURMA performed better than the other two. The values of the *shape index* in case of all the three algorithms were on the higher side than the reference, major reason for this difference being the saw-tooth effect present in the vectorised polygons as no smoothening technique was applied to the segmented results.

4. Conclusion

This paper presented a simple yet reliable and crisp approach for evaluating results of information extraction process on area features using *Omission Error Map, Commission Error Map, completeness*, *correctness* and *quality* parameters. The proposed evaluation approach was also demonstrated for exhaustive evaluation of results of segmentation process to extract building objects as an example. A simple and computationally inexpensive approach was adopted to extract building objects from H-R data. The results of three segmentation algorithms mentioned as BURMA, DBSA, and RSBDNR were compared. Qualitative and quantitative comparisons of the three algorithms were made. A way was also presented to evaluate the extraction results in spatial domain using *Omission Error Map*, *Commission Error Map* and *Overall Quality Map*. BURMA performed best in terms of the proposed evaluation parameters and also in terms of *shape index*, however BURMA did not give flexibility for segmenting a specific range of DN values. RSBDNR was second best in terms of *correctness*, and *quality* but low in terms of *completeness* and *shape index*, however RSBDNR was most flexible at the segmentation stage as it provided option to segment a specific range of pixel values, though vectorisation of the segmented result was not possible in RSBDNR. Quantitatively DBSA showed the second best performance in terms of *completeness* and *shape index*, it also had the facility of vectorizing the segmented result, but qualitatively, and based on *correctness* and *quality*, performance of DBSA was lower as compared to the other two. The spatial visualization of the errors in form of maps gave the exact location of areas under each category and provided an additional appreciation of performance of extraction process at various locations. The use of evaluation parameters namely *completeness*, *correctness* and *quality, Commission Error Map, Omission Error Map and overall quality map* are able to perform an exhaustive quantitative and spatial evaluation of success of segmentation process for extraction of area features. Proposed evaluation parameters along with shape index can be useful for evaluating success of information extraction procedures on area features in all fields dealing with extraction of natural and manmade features such as lakes, glaciers, volcanoes in geological investigations, landslides and dams etc. in disaster studies, and urban features such as buildings, parks in landuse studies, agricultural land, forested areas in agricultural and environmental studies.

References

Ahmadi, S., Zoej, M. J., Ebadi, H., Moghaddam, H. A. and Mohammadzadeh, A., 2010, Automatic Urban Building Boundary Extraction from High Resolution Aerial Images using an Innovative Model of Active Contours. *International Journal of Applied Earth Observation and Geoinformation*, 12, 50–157.

Bacher, U. and Mayer, H., 2005, Automatic Road Extraction from Multispectral High Resolution Satellite images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 29–34.

Blaschke, T., 2010, Object Based Image Analysis for Remote Sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Brekke, C. and Solberg, A. H. S., 2005, Oil Spill Detection by Satellite Remote Sensing. *Remote Sensing of Environment*, 95, 1-13.

Collins, R. T., Hanson, A. R., Riseman, E. M. and Schultz, H., 1995, Automatic Extraction of Buildings and Terrain from Aerial Images. Automatic Extraction of Man-Made Objects from Aerial and Space Images, 169-178.

eCognition, 2004, User Guide 4, Definiens Imaging GmbH, Germany.

ENVI Basic Tools, Sept 2003, User Guide, Research Systems, Inc., USA, 499.

Giada, S., Groeve, T. De., Ehrlich, D. and Soille, P., 2003, Information Extraction from Very High Resolution Satellite Imagery Over Lukole Refugee Camp, Tanzania. *International Journal of Remote Sensing*, 24, 4251-4266.

Holland, D. A., Boyd, D. S., and Marshall, P., 2006, Updating Topographic Mapping in Great Britain using Imagery from High-Resolution Satellite Sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(3), 212–223. doi:10.1016/j.isprsjprs.2006.02.002.

Inglada, J., 2007, Automatic Recognition of Man-Made Objects in High Resolution Optical Remote Sensing Images by SVM Classification of Geometric Image Features. *ISPRS Journal of photogrammetry and Remote Sensing*, 62, 236-248.

Neubert, M. and Meinel, G., 2003, Evaluation of Segmentation Programs for High Resolution Remote Sensing Applications. In Schroeder, M., Jacobsen, K., Heipke, C. (Eds.). *Proceedings Joint ISPRS/EARSeL Workshop High Resolution Mapping from Space 2003*, Hannover, Germany, October 6-8, 2003 (8 p., CD-ROM). http://www.ipi.uni-hannover.de/fileadmin/-institut/pdf/neubert.pdf (accessed May 2008).

Polak, M., Zhang, H. and Pi, M., 2009, An Evaluation Metric for Image Segmentation of Multiple Objects. Image and Vision Computing, 27, 1223-1227.

Udupa, J. K., Leblanc, V. R., Zhuge, Y., Imielinska, C., Schmidt, H., Currie, L. M., Hirsch, B. E. and Woodburn, J., 2006, A Framework for Evaluating Image Segmentation Algorithms. Computerized Medical Imaging and Graphics, 30, 75-88.

Wiedemann, C., Heipke, C., Mayer, H. and Jamet, O., 1998, Empirical Evaluation of Automatically Extracted Road Axes. *Empirical Evaluation Methods in Computer Vision*, 172-187.

Xu, F., Niu, X., and Li, R., 2002, Automatic Recognition of Civil Infrastructure Objects using Hopfield Neural Networks. *ACSM - ASPRS 2002 Annual Conference and FIG 2002 XXII International Congress, Washington DC,* April 19 - 26, 2002.

Zhang, H., Fritts, J. E. and Goldman, S. A., 2008, Image Segmentation Evaluation: A Survey of Unsupervised Methods. Computer Vision and Image Understanding, 110, 260-280.

Zhang, Y. J., 1997, Evaluation and Comparison of Different Segmentation Algorithms. Pattern Recognition Letters, 18, 963-974.