Impervious Surface Detection from Multispectral Images Using Surf

Anu Paulose, Sreeraj M., and Harikrishnan V.

Department of Computer Science,
Federal Institute of Science And Technology,
Cochin, India
anupaulose1990@gmail.com,
sreerajtkzy@gmail.com,
harikrishnan.vadakkath@outlook.com

Abstract. Detection of different regions like impervious surfaces, vegetation and water from a multispectral satellite image is a complex task. This paper introduces a novel idea for impervious surface detection from multispectral images using SURF descriptors. To determine the efficiency of the proposed system, a comparative evaluation is done with other two techniques, namely histogram based and spectral-value-based technique. The result shows that the proposed system outperforms the other two techniques in detecting impervious surfaces like buildings and vehicles with an accuracy of 80.48%. The histogram-based technique and spectral-value-based clustering obtained an accuracy of 61.89% and 68.29% respectively. However, in classifying vegetation the other two techniques outperforms SURF descriptors. The histogram based technique gives an accuracy of 86.46% and an accuracy of 94.35% is obtained by using the spectral-value-based clustering. Whereas SURF based technique gives only an accuracy of 50.71%.

Keywords: Clustering, Histogram, Impervious surface detection, Multispectral, Surf.

1 Introduction

Satellite image processing is a complex task. It is mainly because of the huge amount of data present even in a single image. Object detection is an important research area in satellite image processing. During the past decade, many methodologies have been proposed for automatic identification of objects from multispectral images. It is in fact a type of image classification, where the given image is classified into different object classes like vegetation, water and impervious surfaces (buildings, vehicles etc.,). The recent availability of high-resolution satellite imaging sensors such as IKONOS and QuickBird provide a new data source for impervious surface extraction. In such images the visibility of terrestrial features, especially urban objects, has been increased drastically, which helps in easy detection of small objects like vehicles. Detection of vehicles from

R.C.-H. Hsu and W. Shangguang (Eds.): IOV 2014, LNCS 8662, pp. 237–246, 2014.

satellite images can be used effectively in various fields such as military and surveillance applications to find unauthorzed vehicle entry to a particular area.

In this paper, SURF descriptors are being used for detecting objects, mainly impervious surfaces, from multispectral images. Detection of In order to analyze the efficiency of proposed technique, the detection is also done using Histogram-based method and Spectral-value-based Clustering

This paper is organized as follows: section 2 gives a brief description about the state-of-the-art, section 3 explains the architecture of proposed method, section 4 describes the results and discussions and conclusion is given in section 5.

2 State-of-the-Art

During last four decades, a number of satellite image processing techniques were developed. Based on the literature survey done, a topology of image processing techniques being used was created as shown in Figure 1. The image processing techniques can be grouped into per-pixel, sub-pixel, per-field and object-based approaches. Traditional techniques were on a per-pixel basis, in which, information is extracted from each of the pixels. The spectra of all the pixels are combined to get a spectral signature in such methods. The spectral signature will be having information from all the materials in the pixel which is then used for further processing [1, 2]. However, these methods suffer from mixed pixel problem in which the same pixel may belong to different classes. This will in turn reduce the efficiency of remotely sensed data in per-pixel classifications [3, 4].

The sub-pixel based methods were introduced to avoid the mixed pixel problem in the per-pixel techniques. In these methods, pixels are divided into sub-pixels and features are then extracted from each of the sub-pixels. A fuzzy representation

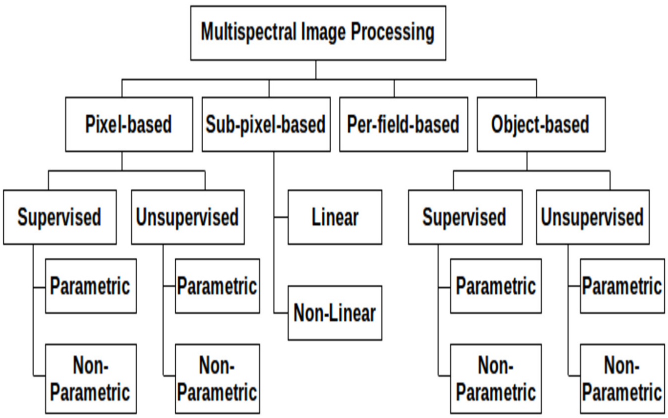


Fig. 1. Topology of Image Classification Techniques

is being used, in which each location contains many partial membership degrees belonging to each of the classes. Sub-pixel classification approaches have been developed to provide a more efficient classification than per-pixel approaches, especially when coarse spatial resolution data are used [3], [5]. Many methods have been introduced for developing a soft classifier such as fuzzy-set theory, softening the output of a hard classification from maximum likelihood, certainty factor [6] and neural networks [7].

Since pixel-based approaches group each of the pixels into a specific class, increase in the spatial frequency may lead to incorrect results. The per-field classifier was developed so as to avoid problems with heterogeneous environments. Many researches have proved the increased classification accuracy of per-field approaches [8–10]. The per-field classifier reduces the noise by using individual units of land called fields [8, 9] . Per-field classification is implemented by integrating both vector and raster data [2] . The vector data are used to subdivide an image into fields, and classification is then conducted based on the fields, thus avoiding intraclass spectral variations. However, per-field classifications are often affected by such factors as the spectral and spatial properties of remotely sensed data, the size and shape of the fields, the definition of field boundaries, and the land cover classes chosen [11] . The per-field classification approach is not very common due to the difficulty in handling both vector and raster data.

An alternate approach is to use an object-based classification [12] , which does not require the use of GIS vector data. Mainly there are two stages in an object-based classification: image segmentation and classification. Using image segmentation pixels are merged into objects, and then a classification is carried out based on objects, instead of individual pixels. This approach has proven to be better when compared to pixel-based approach especially for fine spatial resolution data. The eCognition method is so far the most commonly used object-oriented classification [12].

3 System Architecture

An object based technique is being proposed for identifying different classes in a multispectral image using SURF descriptors. Figure 2 illustrates the generic system architecture. The input images are initially segmented into objects and features are extracted from them. The extracted features are then grouped using clustering algorithm which represents the detected classes. Each of these phases are explained in the following sections.

3.1 Image Segmentation

Segmentation is the process of partitioning an image into regions based on a discontinuity or a similarity criterion. The proposed technique uses Canny Edge Detector for segmenting the images. The working of Canny operator involves multiple stages. Initially, the image is smoothed by Gaussian convolution. Then a simple 2-D first order derivative is applied to highlight the regions of the image

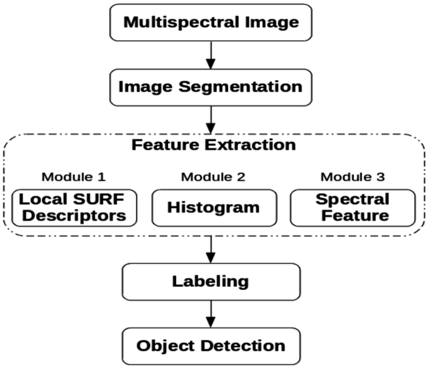


Fig. 2. System Architecture

with high first spatial derivatives. Edges will result in ridges in the gradient magnitude image. The algorithm will then moves along the top of these ridges and all pixels that are not actually on the ridge top are set to zero so as to give a thin line in the output. The tracking process carries out hysteresis with two thresholds: T1 and T2, where T1 < T2. Tracking begins only at a point on the ridge which is higher than T1. Tracking is then continued in both directions until the height of the ridge becomes less than T2. By using this hysteresis, it can be ensured that the noisy edges are not broken up into small edge fragments.

3.2 Feature Extraction

Feature extraction is a special form of dimensionality reduction. It transforms the huge amount of data present in a multispectral image into a reduced representation set of features. Feature extraction is done from each of the segments obtained from the previous phase. The features used in this paper are described in the following subsections.

Speeded-Up Robust Features (SURF). Even though there are many state-of-the-art techniques, most of them uses global features for image classification. Two most popular local feature descriptors in computer vision are SIFT(Scale Invariant Feature Transform) and SURF(Speeded Up Robust Features). They are often used for performing tasks like object recognition. These descriptors are stable under viewpoint and lighting changes, so they are able to cope with significant amounts of image variability. At the same time, discriminative power is achieved by representing feature points as high-dimensional vectors. The technique proposed in [13] extracts SIFT descriptors from the satellite image and uses a graph-cut method for object classification. However, the standard version of SURF is much faster than SIFT. Hence the SURF descriptors are used for classifying the image in this paper. In SURF, the interest points are detected by using Hessian matrix approximation. The use of integral images have drastically reduced the computational complexity. A distribution-based descriptor is used, which makes use of 2D Haar wavelet responses ([14]).

Histogram. The histogram of an image graphically represents the tonal distribution of an image. It plots the number of pixels for each tonal value. The vertical and horizontal axes in the graph represents the tonal variations and number of pixels in that particular tone respectively. The black and dark areas are plotted in left side of the horizontal axis, medium grey in the middle and light and pure white areas towards the right side. Thus, a very dark image creates a histogram having majority of its data points on the left side and center of the graph. Conversely, for a very bright image with few dark areas and/or shadows, the histogram will have most of its data points on the right side and center of the graph.

3.3 Object Detection

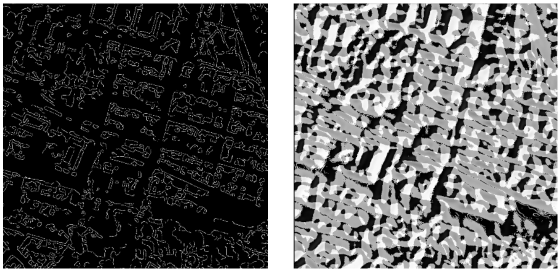
In both SURF-based and spectral-value-based techniques, the extracted features are to be grouped using any clustering algorithms. There are many clustering methods available for classification of a wide variety of data. K-means algorithm is useful for determining the natural spectral regions present in the satellite data. K-means is an unsupervised clustering technique. The user will initiate the algorithm by explicitly specifying the number of clusters to be segmented from the image. In other way, the algorithm can be started in the feature space, each with some pixel clusters defined by its center. By associating each of the pixels to the given nearest centroid, the initial cluster can be created. The mean values of the cluster elements are then computed which replaces the existing centroids. These steps are done repetitively until no more new clusters can be formed.

Whereas in the case of histogram-based technique, the pixels will be automatically grouped into different classes based on their intensity values. Different colors are assigned for the pixels in each class. The impervious surfaces normally appear bright in color. The green colored vegetation areas are seen as red color in a satellite image. Similarly water bodies will be black in color. Keeping these in mind, a rule-based approach was incorporated. For example, green color was assigned to the class which corresponds to red colored pixels.

4 Results and Discussion

4.1 Experimental Setup

Multispectral images of both urban and rural (vegetated) areas were used for the experiments. The dataset used in this paper consists of multispectral images of both urban and rural areas. Two input images used are shown in Figure 4(a) and Figure 5(a). For comparing the performance of image segmentation two edge detectors, namely Sobel and Canny filters, were applied. Results of applying both the filters are given in Figure 3. The result shows that the Canny filter produces better and thin edges compared to the Sobel filter. The Canny edge detector was therefore selected for image segmentation as it outperformed the Sobel filter.



(a) Canny Filter (b) Sobel Filter

Fig. 3. Canny Edge Filter vs. Sobel Edge Filter

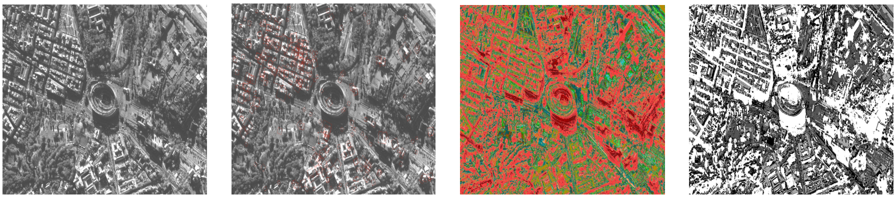
The image segmentation was done using the Canny edge detector for various thresholds (both T1 & T2). When T1 = 1, T2 = 5 the result obtained was very poor, since too many edges in the figure was identified. By increasing the T1 to 10 and the higher threshold to 20, better results were obtained. A more optimal classification was obtained while keeping T1 = 20 & T2 = 40 and the objects were easily detected from the results. When T1 = 30 & T2 = 60, many of the useful edges were eliminated which resulted in poor classification results.

The segmented image is then used as the input to second phase ie., feature extraction. Initially SURF local descriptors are extracted which are to be labelled in the next phase. The second feature being used is the image histogram. The histogram was computed for different number of bins and optimal results were obtained when the number of bins equal to 20. When the histogram was calculated for lesser or higher number of bins, the rate of misclassification increased. Next, the spectral value of each of the pixels is extracted and passed to the next phase.

In the object detection phase, all the extracted features are labelled inorder to detect the various classes present in the image. K-means clustering algorithm is used to group the extracted SURF and spectral features. The spectral-value-based clustering was implemented with the number of clusters being 4, as it gave more accurate results compared to lower or higher number of clusters. The result of object detection using the three techniques for urban area is given in Figure 4 and for rural area is given in Figure 5.

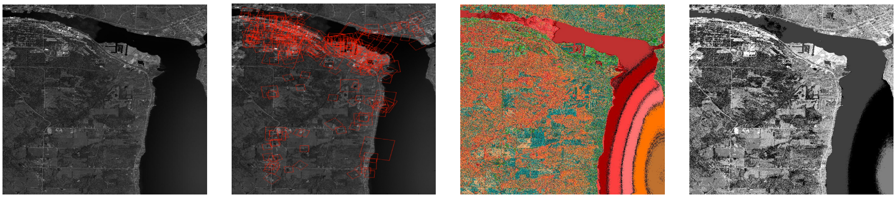
4.2 Performance Evaluation

The proposed technique for object detection, using SURF, proved to be efficient in detecting impervious surfaces, like buildings and vehicles. It outperforms the other two techniques in impervious surface detection. The performance of the proposed system is evaluated using following statistical measures.



(a) Input Urban Image (b) SURF based Object Detection (c) Histogram based Object Detection (d) Spectral value based Clustering

Fig. 4. Object Detection in urban areas



(a) Input Rural Image (b) SURF based Object Detection (c) Histogram based Object Detection (d) Spectral value based Clustering

Fig. 5. Object Detection in rural areas

- Accuracy $= \dfrac{\text{detected number of objects}}{\text{total number of objects}}$
- Sensitivity or True Positive Rate $= \frac{tp}{tp+fn}$
- Specificity or True Negative Rate $= \frac{tn}{tn+fp}$
- Precision or Positive Predictive Value $= \frac{tp}{tp+fp}$
- False Positive Rate $= \frac{fp}{fp+tn}$
- False Discovery Rate $= \frac{fp}{fp+tp}$

where tp = true positive, tn =true negative, fp = false positive and fn = false negative. The confusion matrix developed for Figure 4(a) is shown in Table 1. The image consists of 328 buildings, 475 vehicles, 23 vegetated regions and 10 bare land regions. The confusion matrix for each of the classes is shown in Table 2 which is used for performance evaluation.

After implementing the proposed system using SURF for Figure 4(a), the results obtained are shown in Table 2. Out of the 328 buildings 264 were identified using SURF, which resulted in an accuracy of 80.48%. By using histogram-based technique, only 203 buildings were detected which gave 61.89% accuracy. The spectral-value-based technique produced an accuracy of 68.29%, in which only 224 building were detected. Hence, it is clear that the SURF descriptors outperforms the other three techniques in building detection.

However, other impervious surfaces like vehicles present in the image were not easily identified using the proposed technique. Among the 475 vehicles, only 56 were detected. This is mainly because of the low resolution of images being used in this paper. This problem can be rectified by using high resolution images such as IKONOS. In the case of vegetations, both histogram-based and spectral-value-based techniques outperformed the proposed technique. An accuracy of 94.35% is obtained using the spectral-value-based technique. Histogram-based technique gave an accuracy of 86.46%. Whereas SURF-based technique produced only an accuracy of 50.71%.

Table 1. Generic Confusion Matrix

	Buildings	**Vehicles**	**Trees**	**Bare Land**
Buildings	264	56	7	1
Vehicles	56	419	0	0
Trees	7	0	16	0
Bare Land	1	0	0	9

Table 2. Confusion Matrix For Each Class

	Buildings	**Others**
Buildings	264	64
Others	64	444
Sensitivity = 80.48%, Specificity = 87.40% Precision = 80.48%, FPR = 12.59%, FDR = 19.51%		

	Vehicles	**Others**
Vehicles	56	419
Others	272	289
Sensitivity = 17.07%, Specificity = 40.81% Precision = 11.78%, FPR = 59.18%, FDR = 88.21%		

	Trees	**Others**
Trees	7	16
Others	321	692
Sensitivity = 2.13%, Specificity = 97.74% Precision = 30.43%, FPR = 2.25%, FDR = 69.56%		

5 Conclusion

An impervious surface detection technique from multispectral images was presented. The proposed technique, which uses local SURF descriptors, extracts impervious surfaces with an accuracy of 80.48%. After comparison with two other techniques, namely histogram-based technique and spectral-value-based clustering, results show that our technique was more accurate in detecting impervious surfaces. Among the different impervious surfaces, buildings were more easily identified using the proposed technique than the smaller surfaces like vehicles due to low resolution of images being used. The performance can be improved by the use of high resolution images. However, in rural vegetated areas, SURF descriptors failed to efficiently classify the vegetations. Whereas, both histogram-based and spectral-value-based techniques produced better results in classifying vegetations.

References

1. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing 28(5), 823–870 (2007)
2. Dean, A.M., Smith, G.M.: An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. International Journal of Remote Sensing 24, 2905–2920 (2003)
3. Aplin, P., Atkinson, P.M., Curran, P.J.: Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom. Remote Sensing of Environment 68, 206–216 (1999)
4. Schowengerdt, R.A.: On the estimation of spatial-spectral mixing with classifier likelihood functions. Pattern Recognition Letters 17, 1379–1387 (1996)
5. Cracknell, A.P.: Synergy in remote sensing – what's in a pixel. International Journal of Remote Sensing 19, 2025–2047 (1998)
6. Binaghi, E., Brivio, P.A., Ghezzi, P., Rampini, A.: A fuzzy set accuracy assessment of soft classification. Pattern Recognition Letters 20, 935–948 (1999)
7. Janssen, L.F., Molenaar, M.: Terrain objects, their dynamics and their monitoring by integration of GIS and remote sensing. IEEE Transactions on Geoscience and Remote Sensing 33, 749–758 (1995)
8. Foody, G.M., Cox, D.P.: Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. International Journal of Remote Sensing 15, 619–631 (1994)
9. Foody, G.M.: Image classification with a neural network: from completely-crisp to fully-fuzzy situation. In: Atkinson, P.M., Tate, N.J. (eds.) Advances in Remote Sensing and GIS Analysis, pp. 17–37. John Wiley and Sons, New York (1999)
10. Paola, J.D., Schowengerdt, R.A.: A review and analysis of back propagation neural networks for classification of remotely sensed multispectral imagery. International Journal of Remote Sensing 16, 3033–3058 (1995)
11. Wang, L., Sousa, W.P., Gong, P., Biging, G.S.: Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of panama. Remote Sensing of Environment 91, 432–440 (2004)
12. Fisher, P.: The pixel: a snare and a delusion. International Journal of Remote Sensing 18, 679–685 (1997)
13. Schmitt, D., Mccoy, N.: Object Classification and Localization Using SURF Descriptors. CS 229 Final Project (2011)
14. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. Computer Vision and Image Understanding (CVIU) 110(3), 346–359 (2008)
15. Aplin, P., Atkinson, P.M., Curran, P.J.: Per-field classification of land use using the forthcoming very fine spatial resolution satellite sensors: problems and potential solutions. In: Atkinson, P.M., Tate, N.J. (eds.) Advances in Remote Sensing and GIS Analysis, pp. 219–239. John Wiley and Sons, New York (1999)

16. Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M.: Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing 58, 239–258 (2004)
17. Bloch, I.: Information combination operators for data fusion: a comparative review with classification. IEEE Transactions on Systems, Man, and Cybernetics 26, 52–67 (1996)
18. Foody, G.M.: Status of land cover classification accuracy assessment. Remote Sensing of Environment 80, 185–201 (2002)
19. Sarma, T.H., Viswanath, P., Reddy, B.E.: Single pass kernel k-means clustering method. Sadhan 38, Part 3, 407–419 (2013)
20. Thomas, N., Hendrix, C., Congalton, R.G.: A comparison of urban mapping methods using high-resolution digital imagery. Photogrammetric Engineering and Remote Sensing 69, 963–972 (2003)
21. Walter, V.: Object-based classification of remote sensing data for change detection. ISPRS Journal of Photogrammetry & Remote Sensing 58, 225–238 (2004)