Spectral Clustering Ensemble and Unsupervised Clustering for Land cover Identification in High Spatial Resolution Satellite Images

A Narendranath reddy and NimmagaddaPadmaja

Abstract—Unsupervised clustering plays a dominant role in detailed landcover identification specifically in agricultural and environmental monitoring of high spatial resolution remote sensing images. A method called Approximate Spectral Clustering enables spectral partitioning for big datasets to extract clusters with different characteristic without a parametric model. Various information types are used through advanced similarity criteria. Selection of similarity criterion optimal for the corresponding application is required. To solve this issue a Spectral Clustering Method is proposed which fuses partitioning obtained by distinct similarity representations. This Spectral Clustering Ensemble adopts neural Quantization in the place of Random Sampling, and advanced similarity criterion in the place of Gaussian kernel distance with distinct decaying parameters, and a two level ensemble. The built up areas in the high resolution images can be detected using unsupervised detection. In this process first, a large set of corners from each of the input images are extracted by an improved Harris corner detector. Then, the extracted corners are incorporated into a likelihood function to discover candidate regions in each input image. Given a set of candidate build-up regions, in the second stage, the problem of build-up area detection is concised as an unsupervised grouping problem. The performance of these algorithms is evaluated by Accuracy, Adjusted Rand Index (ARI) and Normalized Mutual Information (NMI). Experimental results show a significant betterment of the resulting partitioning obtained by the proposed ensemble, with respect to the evaluation measures in the applications.

Index Terms— Spectral clustering (SC), cluster ensemble, clustering methods, geodesic similarity, land-cover identification, pattern identification, similarity criterion.

1 INTRODUCTION

The Spatial resolution of remote-sensing images has become below meter resolution with increase in number of spectral bands. Because of very high spatial resolution enables detailed spectral, textural, or structural representation to get accurate information in either supervised or unsupervised manner for agricultural and environmental control, urban monitoring and surveillance, disaster management, and homeland security (see [1]–[3] and references therein). Supervised approaches require more expert time and labelled training points (which are often hard and expensive to get since they can be purely obtained only with staticmeasurements [4]). It becomes more challenging for agricultural monitoring where crop calendar, seasonal changes, and regional changes play a significant role in spectral signatures of the land cover/use classes [5], [6]. Consequently, achieving supervised classification performances with unsupervised approaches (which need no labelled points and only limited expert interaction) is of great importance for effective agricultural control at the state level.

Popular unsupervised k-means method and its variant ISODATA methods are usually poor in performance due to their centroid based parametric methods which produce (hyper) spherical clusters [5]. Approaches alternatively without centroid models (hierarchical clustering [8], neural clustering [9], spectral clustering (SC) to name a few) are proved more successful than k-means for remote-sensing images. Among them, SC, which has the ability to extract clusters of distinct characteristics without any parametric model, has become recently popular [10]–[13]. Because of its computational complexity, SC is performed for very high spatial resolution remote-sensing images through a two-step approach, i.e., the Spectral Clustering is applied on the data representatives chosen by sampling or quantization [10], [16]. This enables the usefulness of Spectral Clustering for larger remote-sensing images and combines different information types like distance, density, and topology to produce a powerful

information representation for extracting accurate cluster structure [17]. Still, it requires an optimum selection for similarity definition adjust to the application requirements. An alternative of empirical determination of optimal conditions, the partitioning acquired with different settings can be fused by ensemble learning.

The ensemble methods consolidate the partitioning obtained by distinct input or feature sets, distinct methods, or the same method with many parameter settings, using various ways such as majority voting, evidence accumulation, hyper graph operations, metaclustering, or mixture models [10], [18]. When they are used to merge the decisions of the same method with several parameter settings, as in Spectral Clustering ensembles (SCE) [10], they avoid the need to determine the optimal parameter values. For instance, the SCE in [10], which uses random sampling (Nystrom method [14]) in conjunction with maximum voting and metaclustering algorithm, clubs partitionings obtained with distinct kernel parameter values in similarity definition, for segmentation of relatively small SAR images [10]. The same approach is used for image segmentation [22] also. In spite of the ensemble approach, distinct parameter windows, specific to the datasets, are used to attain high performance [10]. In addition, when random sampling is used for Spectral Clusterimg of large datasets, out-of-sample labelling may become problematic [16], [24].

In this paper, a Spectral Clustering Ensemble (SCE) is proposed for unsupervised clustering of large remote-sensing images. The SCE is novel in three ways. First, we acquire the data representatives using neural gas [25] (instead of Nystrom approximation), to get less quantization errors and address out-of-sample labelling problem. Second, we employ the recent similarity definitions [13], [17] utilizing various information types, as a substitute to the ordinary Gauss kernel distance with different parameter values. Third, we have a two-level ensemble method which acquires a combined decision for each similarity criterion and then fuses these decisions into a unique label. We test the SCE performance on commonly available remote-sensing data and remote-sensing images for agricultural control. Depending on the calculation by Accuracy, adjusted Rand index (ARI) and normalized mutual information (NMI), we show that the proposed SCE achieves high performance with respect to all three evaluation measures.

2 IMPLEMENTATIONOFCLUSTERING PROCESS

2. 1 Spectral Clustering Ensemble

The proposed SCE combines the partitioning obtained by Spectral Clustering with different similarity criteria. It has two sections: Spectral Clustering and ensemble learning. Now, this Clustering applies SC on a reduced set of data representatives preferred by sampling or quantization [12], [14]–[16] so that it enables the use of SC in remote-sensing images to obtain clusters of different characteristics without using a parametric model. Sampling based SCs are faster than quantization based SCs at a cost of a lower clustering performance because of relatively high quantization errors and out-of-sample labelling problem [12], [16], [24]. The neural gas quantization [25] is shown more successful than k-means and its variants for this SC in remote-sensing image analysis [12], [13]. After choosing representatives, this Spectral Clustering is similar to the normal Spectral Clustering approaches. Still, the use of representatives in proposed SC enables multifold based similarity definitions like data topology and local density, in addition to the common distance-based similarities [13], which yields distinct partitioning. Ensemble learning combines the emerging partitionings to have a uniqueness on the label of data representatives. Then, the label of each representative is referred to its corresponding image pixels. An outline of the SCE method is shown in Fig. 1 and summarized in Table I. We now briefly explain the SCE.

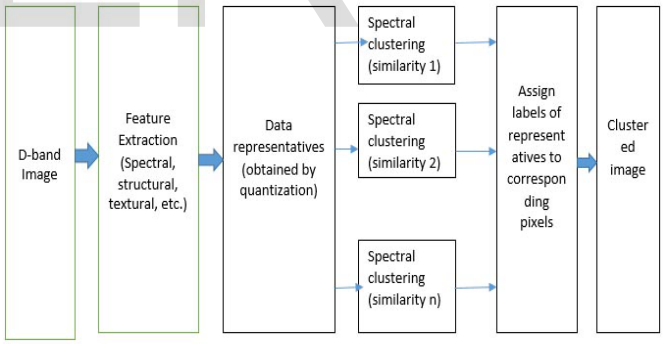


Figure. 1. ASC ensemble (ASCE). Spectral or spatial features relevant to remote-sensing application are extracted. The *Nr*data representatives are selected by neural gas quantization of these features. The data representatives are partitioned by SC using *ns* similarity criteria in Section II-C with n_{km} k-means runs. The resulting partitions are merged with cluster ensemble as explained in Section II-D. The ensemble labels of the representatives are assigned to their corresponding data points (pixels).

TABLE 1
Approximate Spectral Clustering Ensemble

1. Spectral Clustering Ensemble

Input: $n_s n_{km}$ partitioning of the N_r representatives.
-Obtain a similarity matrix S_{CEa} using the number of identically laballed representatives among n_{km} different partitionings of each similarity criterion.
-Apply spectral clustering based on these S_{CEa} to obtain n_s first level ensemble partitionings.
-Obtain a similarity matrix S_{CEb} using the number of identically labelled representatives among the arising n_s ensemble partionings.
-Apply spectral Clustering based on S_{CEb}.
Output: Final partitioning of N_r representatives.
2. Assign the representatives' labels to their corresponding pixels.
==========

2. 1. 1 Data Representatives by Neural Gas Quantization

As a first step in ASCE, we obtain data representatives by neural gas quantization [25]. Being a neural learning paradigm, the neural gas [25] quantizes the data points in a topology preserving manner based on neighborhood ranking. Its randomly initialized neural units evolve to the quantization representatives with an iterative learning: a data point v is randomly selected from the dataset at each step, and its best-matching unit (BMU) w_i is found by the minimum distance

$$||v - wi|| \leq ||v - wj||. \quad (1)$$

Then, the neural units w_j neighbor to the BMU w_i are determined by a neighborhood function $h_\tau(w_j) = \exp(-\rho_{wi}/\tau)$

based on ρ_{wj} which is the rank of the distance of w_j to v. Note that the rank $\rho_{wi} = 0$. Finally, w_i and its neighbors w_j are adapted by

$$w_j(t+1) = w_j(t) + \alpha(t)h_\tau(w_j)(v - w_j(t)) \quad (2)$$

Where $\alpha(t)$ is a learning parameter decaying by time. When this iterative process is completed by a predetermined stop criterion, the neural units become the quantization representatives, achieving a lower quantization error than k-means [25].

2. 1 2 SC of Data Representatives

SC methods are multiple learning approaches, which successfully extract clusters with distinct characteristics, these methods follow nonparametric model approach and easy implementation [28]–[30]. They are based on eigen value decompositon of a graph Laplacian derived from pairwise similarities of the data points. This results in heavy computational load, which makes them unusable for large datasets such as remote sensing images having high spatial resolution. To exutilize SC advantages for these datasets, ASC performs SC on a reduced set of data representatives.

Being associated with relaxed optimization of graph-cut problems, SC constructs a graph *Laplacian* matrix L based on some optimization basis [28]–[30]. Due to the fact that there is no clear advantage among SC methods as long as a normalized graph Laplacian is used [31], [32], we employ the SC method in [29] to cluster the data representatives obtained by neural gas: Let $G = (V, S)$ be a weighted undirected graph where its nodes V represent the representatives and the edges S are the pairwise similarities between them. The normalized Laplacian matrix L_{norm} is defined as

$$L_{norm} = D^{-1/2}SD^{-1/2} \quad (3)$$

based on a similarity matrix S and its diagonal degree matrix with $d_i = \sum_j s(i,j)$. Then, the SC method is as follows.

1) Construct a similarity matrix S showing the pairwise similarities of the N_r representatives to be clustered;
2) Calculate D and L_{norm} using the similarity matrix S;
3) Find the k eigenvectors $\{e_1, e_2, ..., e_k\}$ of L_{norm}, associated with the k highest eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_k\}$;
4) Construct the $N_r \times k$ matrix $E = [e_1 e_2 ... e_k]$ and obtain $N_r \times k$ matrix U by normalizing the rows of E to have norm1, i.e.,

$$u_{ij} = \frac{e_{ij}}{\sqrt{\sum_k e_{ik}^2}}$$

Cluster the N_r rows of U with k-means into k clusters.

The eigendecomposition of the graph Laplacian (ideally) produces a data projection where submanifolds (clusters) are well separated. A simple method in Step 5, such as k-means, is hence expected to produce a clear delineation among clusters. However, this is often not possible due to complex data structures, resulting in intersecting submanifolds. Therefore, consecutive k-means runs may result in distinct partitionings, due to the randomness in k-means algorithm [13].

2. 1. 3 Similarity Criteria for ASC

The similarity criterion to be set for S plays a significant role to achieve an accurate cluster extraction. The pairwise similarities in SC $s(i,j)$s are traditionally determined by a Gaussian kernel based on the (Euclidean) distances

$d_{Euc}(x_i,x_j)$, with a global decaying parameter σ (to be optimally found through experiments [29] or a local $\sigma_i$reflecting the distance to the kth-nearest neighbor of x_i [33])

$$S_{Euc}(i,j) = \exp\left(-\frac{d_{Euc}(x_i x_j)}{2\sigma_i\sigma_j}\right) \quad (4)$$

However, for ASC, new information types such as topology, density can be embedded into S for more effective definitionfor pairwise similarities of the data representatives [12], [17], [34]. A recent approach [12] uses a similarity measure (CONN) that exploits local density together with data topology on the representative level. $CONN(i,j)$ shows the number of data points for which the representatives $w_i$and $w_j$are the pair of the bestmatching and the second-BMUs. In other words, $CONN(i,j)$ represents the local density distribution inside the subregions$V_{ij}\cup V_{ji}$of the Voronoi polygons V_i and V_j, (V_i is the set of data points v for which $w_i$is the closest representative), i.e.,

$$CONN(i,j) = |V_{ij}\cup V_{ji}| \quad (5)$$

where the sub-Voronoi polyhedron $V_{ij}$is

$$V_{ij} = \{v \in V_i : ||v - w_j|| \leq ||v - w_k|| \ \forall k \neq i\} \quad (6)$$

By its definition, CONN is a weighted version of the induced Delaunay triangulation in [25] showing the neighbors according to the manifold, where weights indicate how the data points are distributed within the Voronoi polygons. CONN thus produces more accurate partitionings than those obtained by distance based approaches [12], [35]. The distance information is also integrated with CONN for ASC by a hybrid similarity criterion S_{hyb}[34] where

$$S_{hyb}(i,j) = S_{Euc}(i,j) \times exp\left(\frac{CONN(i,j)}{max_{i,j} CONN(i,j)}\right) \quad (7)$$

The hybrid S_{hyb}actually scales the distance-based similarity with respect to local density distribution and data topology. If $CONN(i,j) = 0$, then two representatives $w_i$and $w_j$are not neighbors, which results in $s_{hyb}(i,j) = s_{Euc}(i,j)$. Otherwise, their similarity is scaled by $[1,e]$, producing a greater similarity upto a scale of e for the maximum $CONN(i,j)$.

Recently, geodesic-based approaches are also proposed for ASC [13], [17]. To calculate geodesic distances, a preliminary step is to determine a neighborhood graph showing the neighbor representatives in the manifold. A traditional way toconstruct this graph is the use of (mutual) k-nearest neighbor ($k-nn$) graph. If $w_i$and $w_j$are among their k-closest neighbors, they are neighbors. Then, the geodesic distance between $w_i$and $w_j$is the sum of Euclidean distances of their shortest path

$$d_{geoknn}(w_i,w_j) = \sum_{lm \in SP_{knn}} d_{Euc}(l,m) \quad (8)$$

where$SP_{knn}(w_i,w_j)$ is the set of edges in the shortest path between $w_i$and $w_j$calculated with $d_{Euc}$and $k-nn$graph. The parameter k should be optimally set; however, k may be different for each representative. As a topology-based alternative reflecting local characteristics, CONN can be used for neighborhood graph in calculating geodesic distances to find specific number of neighbors for each representative [17]. The geodesic distance d_{geoadj}based on CONN graph using Euclidean distances $d_{Euc}$is calculated as

$$d_{geoadj}(w_i w_j) = \sum_{lm \in SP_{adj}(w_i w_j)} d_{Euc}(l,m) \quad (9)$$

where$SP_{adj}(w_i,w_j)$ is the set of edges in the shortest path between $w_i$and $w_j$based on $d_{Euc}$and CONN neighborhood graph.

Instead of distances, local density distribution can also be used for geodesic distance calculation. Namely, using density-based distance

$$d_{CONN}(w_i,w_j) = \begin{cases} e^{-\frac{CONN(i,j)}{max_{y,z} CONN(y,z)}}, if\ CONN(i,j)>0 \\ \infty, \quad Otherwise \end{cases} \quad (10)$$

a geodesic distance including both data topology and the data distribution can be defined as [17]

$$d_{geoconn}(w_i,w_j) = \sum_{lm \in SP_{conn}(w_i,w_j)} d_{CONN}(l,m) \quad (11)$$

$SP_{conn}(w_i,w_j)$ is now the set of edges in the shortest path between $w_i$and $w_j$with respect to d_{CONN}distance and CONN neighborhood. In addition, to exploit all available information for ASC on the representative level, a hybrid approach $d_{geohyb}(w_i,w_j)$ merging both distance and density using CONN graph can be defined as

$$d_{geohyb}(w_i,w_j) = \sum_{lm \in SP_{hyb}} d_{Euc}(l,m) d_{CONN}(l,m) \quad (12)$$

The geodesic distance-based similarities are then obtained by replacing d_{Euc}in (4) with the corresponding distance criterion. They achieve better clustering accuracies for a wide variety of datasets with different clustering statistics [13].

2. 1. 4 Cluster Ensemble

In order to achieve better performances with ensembles than with single partitionings, it is important to have diverse results obtained by different feature sets, subsets, clustering

methods, or different parameters for the same method. In our approach, we produce a set of (neural gas) data representatives and obtain diversity with various similarity criteria (measuring distance, topology, and density) and random initializations for k-means step in SC. By the use of data representatives, graph-based ensemble approach [18]—which is infeasible for large data— can be easily applicable for an effective decision fusion. A graph $G = (V, S_{CE})$ is constructed based on $n_s \times n_{km}$ partitionings obtained by n_s similarity criteria and n_{km} k-means runs per similarity criterion, where $S_{CE}(i,j)$ shows the pairwise similarity of the representatives w_i and w_j with respect to the number of partitionings for which they have the same label, i.e.,

$$S_{CE}(i,j) = \sum_{m=1}^{n_s} \sum_{k=1}^{n_{km}} S_{m,k}(i,j) \qquad (13)$$

where $S_{m,k}(i,j) = 1$ if w_i and w_j are in the same cluster; otherwise, $S_{m,k}(i,j) = 0$. An SC is applied to this graph to get the consensus labels for the representatives. This one-level ensemble is called as ASCE1.

Alternative to the traditional approach of fusing all results into one as in ASCE1, a two-level ensemble process (ASCE2) is used. In ASCE2, first get an ensemble for a similarity criterion s_* from its corresponding partitionings obtained by n_{km} k-means runs, based on a similarity matrix $S_{CE2a} = \sum_{k=1}^{n_{km}} S_{s*,k}$. Then fuse n_s ensemble results of each similarity using $S_{CE2b} = \sum_{m=1}^{n_s} S_m$.

By its construction, the two-level ASCE2 first addresses the randomness in k-means and then exploits distinct results obtained by different information types. Note that fuse the clustering labels at the representative level and then determine the labels of the data points according to the ensemble labels of their representatives. Table I outlines the proposed ASCE.

2.2 Automated Extraction of Candidate Built-Up Regions

Given a set of high-resolution remote sensing images covering different scenes, our objective is to simultaneously detect built-up regions from them. Generally speaking, urban environment is replete with corners from building roofs, road marks, and other man-made objects. If we could detect all such corner points from images, the built-up regions would be naturally implied from the density of corners. Thus, in this section, we use the corner feature to infer the locations of potential built-up regions in the given images. Corner detection has been a longstanding problem in computer vision. In the literature, a large number of methods have been proposed, and the most famous one should be Harris corner detector [7]. In order to achieve a reliable extraction of corners from built-up areas, we proposed two criterions, which take both local and global constraints into consideration, to validate and filter a large set of initial extracted Harris corners.

Extraction of Candidate Built-Up Regions by Grouping Corner Points

Generally, corners in the built-up area tend to closely locate in the neighboring spatial domain with high density. For a non-built-up area, these corners are more likely to be sparsely distributed.

This means that, if an image pixel (x_i, y_i) belongs to a built-up area, we expect that there are more corner points in its neighborhood. With this observation, we define the following likelihood function to measure the possibility that a pixel (x_i, y_i) belongs to a built-up area:

$$LS(x_i, y_i) = \sum_{k=1}^{N_2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}}{2}\right) \qquad (14)$$

where (x_k, y_k) represents the spatial coordinate of the extracted N_2 corner points, for $k = 1,\ldots,N_2$.

The likelihood function highlights the built-up region in the pixel neighborhood. If it is a good candidate for the built-up region, a high value $LS(x_i, y_i)$ is expected. On the basis of this, we first compute the likelihood of each pixel (x_i, y_i) and then obtain candidate built-up regions by the binarization processing of an image pixel based on the threshold value, we provide the detected candidate built-up area (as a yellow curve) from Fig. 2 in Fig. 3(a). We also provide the ground truth data in Fig. 3(b).

Feature Extraction of Candidate Built-Up Regions

Since the built-up region has unique texture in comparison with a natural area, here, we use the texture feature to describe the built-up region. In our work, the texture feature is obtained by the following three steps.

Step 1) Convolving the panchromatic band of each candidate region $R_i$with a Gabor filter bank [11] at K scales and S orientations (in the experiment, we set K = 4 and S = 6) to generate filter responses.

Step 2) The resultant filter responses are aggregated and clustered into textons using the k-means algorithm, which are further used to form a texton dictionary Ω.

Step 3) Ak-dimensional histogram descriptor h_i is constructed for each candidate region $R_i$by labeling each filter response with the texton which lies closest to it in the dictionary.

3 EXPERIMENTAL RESULTS

We evaluate the performance of the proposed ASCE2 performance based on accuracy (percentage of correctly labeled data), Adjusted Rand Index (ARI) [36], and Normalized Mutual Information (NMI) [37] using five remote-sensing images from www.satimagingcorp.com (Geoeye-1, Worldview-2, Ikonos, Spot-7) which are commonly available.

3. 1 Datasets

The Bondi beach, Autralia (September, 2010) image is taken from Geoeye-1satellite is available from www.satimagingcorp.com . It is a remote-sensing image of 3000 × 3000 pixels. It has six classes: beach, ocean, road, park, residential, and industrial. Fig. 2 (a) shows Natural color composite using red, green, and blue spectral bands of the Geoeye-1 satellite image. Then the pixels are labeled to these eight classes to represent the ground truth.

As a high spatial resolution example, we use a geoeye-1 image (0.5-m spatial resolution) for land-cover identification. The (3000 × 3000-pixel) image covers an approximately 3.5km² region on Copacabana beach with different classes like Sea, Road, forest, Residential area, Beach.

Fig. 3 (a) provides a natural color composite of the image. The test pixels were randomly determined proportional to their land coverage in the study area, using field study and domain knowledge of the national experts [39]. Supervised classification accuracies for these classes were 76.4% for eight multispectral bands and 83.2% for fused features (eight multispectral and four Gabor features). We therefore use fused features to evaluate the proposed ASCE2. Fig. 3 (b) Provides clustermap with different classes.

The image from Ikonos Satellite is taken to determine the type of land cover. Fig. 4. (a) Shows the natural color composites of Rio-de-Janeiro port, Brazil for evaluation of land cover. It is a remote-sensing image of 1950 × 1950 pixels. It has classes like ocean, road, parks, residential and industrial areas. Fig. 4. (b) Shows the clustering map of the image.

Fig. 5. (a) Shows the natural color composites of Rice fields in Inkadate city of Japan. It is a remote sensing image of 3000 × 3000 pixels taken by Geoeye-1 Satellite on July 15, 2010. This is to observe the land cover of Rice fields around the City. This image contains several classes like Rice fields, Roads, Residential area etc. By using the clustering ensemble on this image we obtained the clustering map of the image.

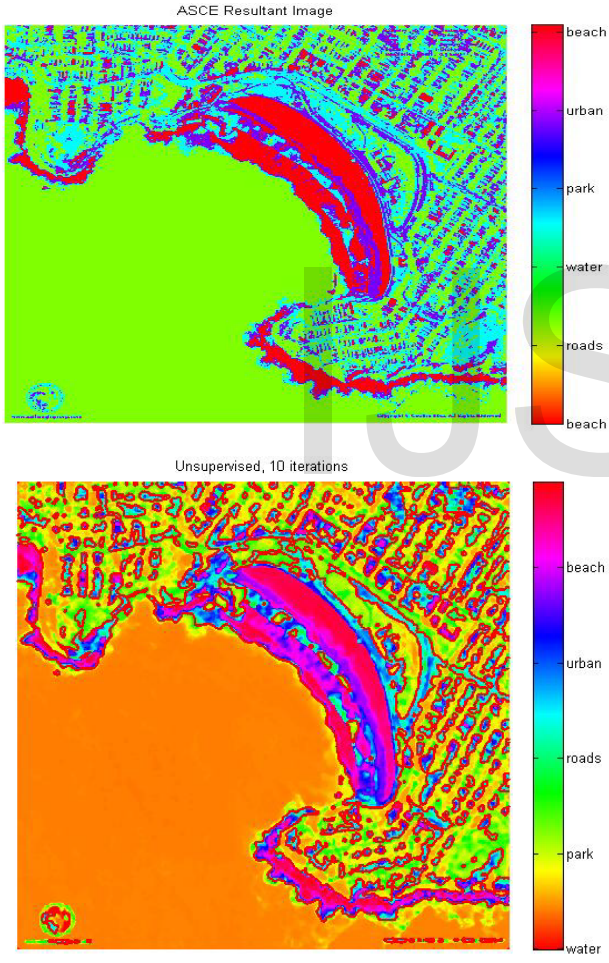


Figure. 2. Bondi beach, Australia from Geoeye-1. (a) Natural color composite of the dataset (3000×3000 pixels). (b) Resulting clustermap obtained by ASCE. Each color indicates one of the six extracted clusters. (c) Clustermap obtained by unsupervised clustering.

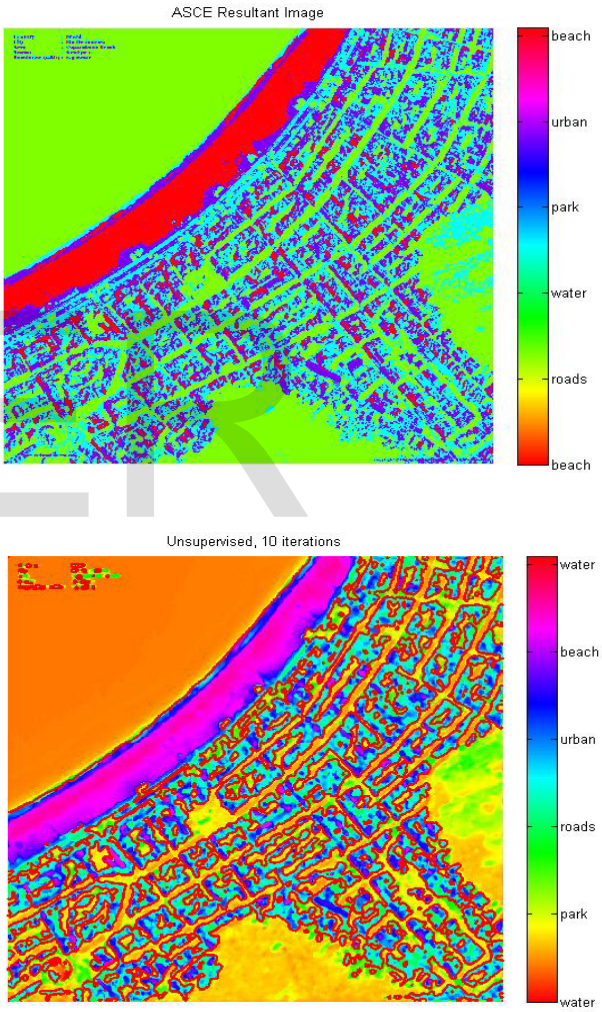
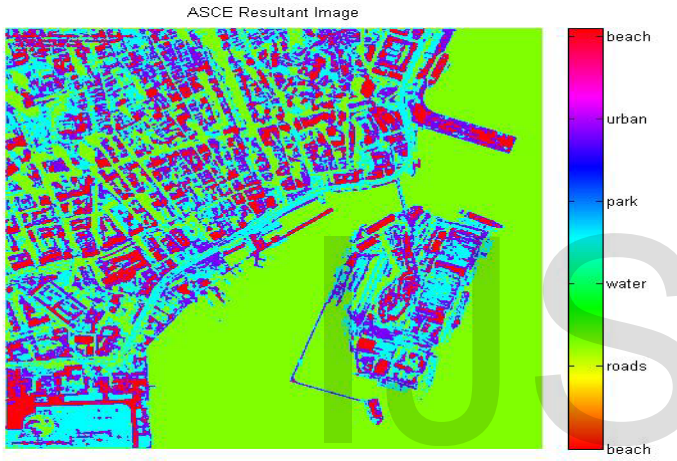
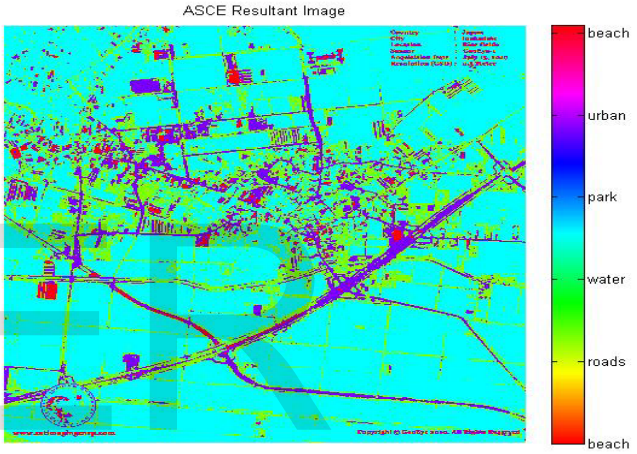
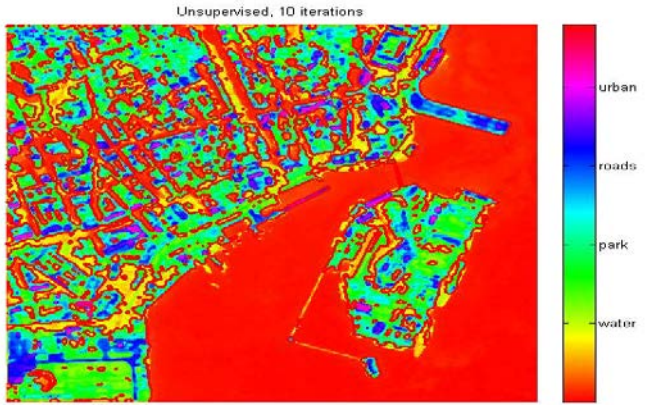


Figure. 3. Geoeye-1 image of Copacabana beach. (a) Natural color composite using red, green, and blue spectral bands of the Worldview2 image. (b) ASCE2 cluster map with land cover types. (c) Clustermap obtained by unsupervised clustering.







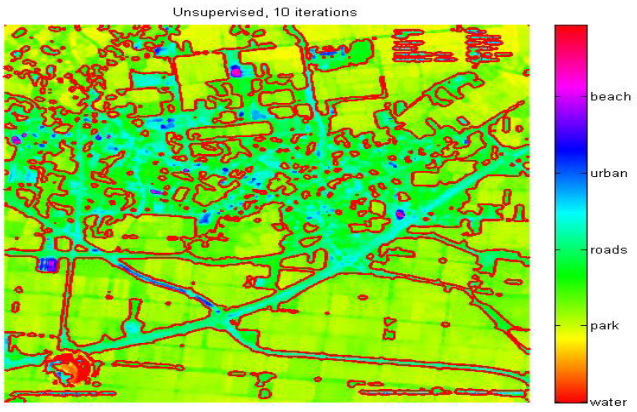


Figure. 4. Geoeye-1 image of Rio-de-janeiro. (a) Natural color composite using red, green, and blue spectral bands of the Worldview2 image. (b) ASCE2 cluster map with land cover types. (c) Clustermap obtained by unsupervised clustering.

Figure. 5. MultitemporalRapidEye images of (a) Rice fields in japan (b) ASCE2 cluster map with land cover types. (c) Clustermap obtained by unsupervised clustering.

3. 2 Performance Evaluation

We use three measures to evaluate the partitionings of these remote-sensing images: accuracy, ARI, and NMI. Accuracy is the percentage of correctly clustered data points. ARI [36] is a measure of agreement between the labels gathered by a clustering process and the other labels defined by external criteria for the same data [40]. ARI recognize not only the correct separation of data points into different clusters but also the relation between data points of the same cluster, to sensitively classify the relation between each datum and its target label for multiclass problems [40]. NMI is used to compare the resulting partitionings in an information theoretical way [37].

TABLE 2

Percentage of Land Cover Detected Using Spectral Clustering Ensemble

Label	SCE			
	Gondi beach	**Copacabana beach**	**Rio-de-janeiro port**	**Rice fields Japan**
Park	96.98	116.77	95.18	113.45
Water	110.82	90.50	94.73	89.10
Beach	94.17	93.32	102.04	96.51
Urban	92.05	105.15	98.64	108.17
Roads	105.73	89.98	121.01	93.88

Confusion Matrix:

- True positive = correctly identified
- False positive = incorrectly identified
- True negative = correctly rejected
- False negative = incorrectly rejected

- True Positive Rate (TPR): It is the proportion of positive cases that were correctly identified

$$TPR = \frac{A}{\mathrm{A + B}}$$

- Accuracy (AC): is the proportion of the total number of predictions that were correct.

$$Accuracy = \frac{A+D}{A+B+C+D}$$

- The False Positive Rate (FPR): is the proportion of negatives cases that were incorrectly classified as positive.

$$FPR = \frac{C}{C+D}$$

Table 2 and table 3 show the percentage of land cover that has been detected correctly by Spectral Clustering Ensemble algorithm and Unsupervised Clustering algorithm using Harris Corner Detector respectively. Apart from the classes mentioned in the table, there is chance of other areas. We have considered only the areas which are common in the different images taken by different Satellites.

TABLE 3

Percentage of Land Cover Detected Using Unsupervised Clustering

Label	Unsupervised			
	Gondi beach	**Copacabana beach**	**Rio-de-janeiro port**	**Rice fields Japan**
Park	96.07	102.56	105.19	117.94
Water	113.67	94.34	96.28	127.67
Beach	90.35	86.33	89.51	81.53
Urban	93.31	94.77	94.96	87.00
Roads	103.33	110.93	106.22	94.31

TABLE 4

Performance Measures of Spectral Clustering Ensemble

Method	SCE			
Area	Gondi beach	Copacabana beach	Rio-de-janeiro port	Rice fields Japan
TPR	0.8400	0.8557	0.8630	0.8582
FPR	0.1600	0.1443	0.1370	0.1418
Accuracy	89.9927	90.9783	91.4297	91.1360
ARI	0.6390	0.6646	0.6860	0.6672
NMI	0.1141	0.0413	0.5976	0.0613

TABLE 5

Performance Measures of Unsupervised Clustering

Method	Unsepervised			
Area	Gondi beach	Copacabana beach	Rio-de-janeiro port	Rice fields Japan
TPR	0.8403	0.8544	0.8612	0.8565
FPR	0.1597	0.1456	0.1388	0.1435
Accuracy	90.0094	91.8940	92.3246	92.0380
ARI	0.6218	0.6194	0.6986	0.6911
NMI	0.0824	0.0587	0.0663	0.0740

Observing the values of True Positive Rate (TPR) and False Positive Rate (FPR) from tables 4 and 5, suppose if TPR is 0.84 or 84%, it means that every time you call a positive there is a probability of 16% of that being wrong, which is the FPR. Similarly if FPR is 16% or 0.1600, which means that every time you call a negative there is probability of 84% being right.

By analyzing each pair of elements ARI will measure not only the correct separation of elements belonging to different classes but also the relation between elements of the same class. In a certain way this measure pays more attention to the relation between elements than to the relation between each element and its target label. We can say that ARI evaluates the capability of the algorithm to separate the elements belonging to different classes. From table 4 and table 5, the values for Adjusted Rand Index (ARI) are closer to 1 than 0. This indicates that these algorithms will separate the elements belonging to different classes effectively. If the ARI value is 1 (maximum) meaning that the algorithm is doing the exact distinction between classes. Observing another parameter Normalized Mutual Information (NMI) from table 4 and table 5, the value is very close to 0, which indicates that only very small portion of two clusters are similar. That means cluster are almost unique. Accuracy is another important factor which is the proportion of the total number of predictions that were correct. From table 4 and table 5, it is clear that large portions of the original image is predicted correctly by using Spectral Clustering Ensemble and Unsupervised clustering using Harris Corner Detector algorithms.

4 CONCLUSION

High spatial resolution has brought new test for remote-sensing image analysis while creating emerging applications due to improved abilities of fine details for accurate agriculture, homeland security, and urban monitoring. A longstanding task for remote-sensing images, and for those with high spatial resolution, is the necessity of accuratelabelled training sets (which can only be obtained purely from the field visits) to achieve effective land-cover maps with many supervised approaches [41].

While active learning approaches [42] address this by reducing the required labeled samples and guiding the user to label few selected samples, unsupervised clustering uses no labeled samples. Recently, ASC methods enable the spectral partitioning of large datasets such as remote-sensing images with high spatial resolution, to utilize its advantages of extracting clusters with various characteristics without a parametric model. The ASC methods also enable effective manifold-based information representation criteria that are optimally selected for corresponding application. To harness different information types in ASC without a need for optimality selection, we introduced an SCE composed of three stages: 1) neural gas quantization to obtain data representatives; 2) SC of these representatives with recent similarity criteria derived from different information types exploiting distinct manifold characteristics; and 3) a two-level graphbased ensemble of obtained partitionings. We showed that the SCE improves

the clustering quality with respect to three measures (accuracy, ARI, and NMI) for five remote-sensing images. Based on its success on remote-sensing applications for land-cover identification, the SCE can be used for accurate cluster extraction from large remote-sensing-images in an unsupervised manner. The SCE can be helpful especially for agricultural monitoring (encompassing a whole country) where it is often hard and expensive to obtain labeled training samples for supervised classification methods.

The SCE is based on data representative level ensemble approach, which in turn limits the ensemble on one kind of representative selection (neural gas in our study). However, the use of different quantization approaches (such as k-means, selforganising maps, or their variants) or different initializations may produce distinct partitionings, which are then to be fused by an ensemble approach. Yet, graph-based ensemble would be infeasible at the data level for large remote-sensing images due to its complexity, while maximum voting would be unsuccessful based on our preliminary studies. An ensemble approach tailored for ASC of large datasets would help for more effective partitionings. With its current settings using neural gas quantization, the proposed SCE already achieves significant performances for the remote-sensing images in this study.

REFERENCES

[1] L. A. Ruiz, A. Fdez-Sarria, and J. A. Recio, "Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study," in *Proc. Int. Arch. Photogramm. Remote Sens*., 2004, vol. 35, pp. 1682–1750.

[2] E. M. Wood, A. M. Pidgeon, V. C. Radeloff, and N. S. Keuler, "Image texture as a remotely sensed measure of vegetation structure," *Remote Sens. Environ.*, vol. 121, pp. 516–526, Jun. 2012.

[3] J. Yuan, D. Wang, and R. Li, "Remote sensing image segmentation by combining spectral and texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 16–24, Jan. 2014.

[4] P. Mitra, U. Shankar, and S. Pal, "Segmentation of multispectral remote sensing images using active support vector machines," *Pattern Recognit. Lett.*, vol. 25, no. 9, pp. 1067–1074, Jul. 2004.

[5] R. Xu and D. Wunsch II, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 645–678, May 2005.

[6] K. Tas¸demir, P. Milenov, and B. Tapsall, "A hybrid method combining SOM-based clustering and object-based analysis for identifying land in good agricultural condition," *Comput. Electron. Agric.*, vol. 83, pp. 92– 101, 2012.

[7] M. L. Goncalves, M. L. A. Netto, J. A. F. Costa, and J. K. Zullo, "An unsupervised method of classifying remotely sensed images using Kohonen self-organizing maps and agglomerative hierarchical clustering methods," *Int. J. Remote Sens.*, vol. 29, no. 11, pp. 3171–3207, 2008.

[8] K. Tas¸demir, P. Milenov, and B. Tapsall, "Topology-based hierarchical clustering of self-organizing maps," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 474–485, Feb. 2011.

[9] T. Villmann, E. Merényi, and B. Hammer, "Neural maps in remote sensing image analysis," *Neural Networks*, vols. 3–4, no. 16, pp. 389–403, 2003.

[10] X. Zhang, L. Jiao, F. Liu, L. Bo, and M. Gong, "Spectral clustering ensemble applied to SAR image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2126–2136, Jul. 2008.

[11] K. Ersahin, I. G. Cumming, and R. K. Ward, "Segmentation and classification of polarimetric SAR data using spectral graph partitioning," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp. 164–174, Jan. 2010.

[12] K. Tas¸demir, "Vector quantization based approximate spectral clustering of large datasets," *Pattern Recognit.*, vol. 45, no. 8, pp. 3034–3044, 2012.

[13] K. Tas¸demir, B. Yalcin, and I. Yildirim, "Approximate spectral clustering with utilized similarity information using geodesic based hybrid distance measures," *Pattern Recognit.*, vol. 48, no. 4, pp. 1459–1471, 2015.

[14] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nyström method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 214–225, Feb. 2004.

[15] L. Wang, C. Leckie, K. Ramamohanarao, and J. C. Bezdek, "Approximate spectral clustering." in *PAKDD*, vol. 5476, T. Theeramunkong, B. Kijsirikul, N. Cercone, and T. B. Ho, Eds. New York, NY, USA: Springer, 2009, pp. 134–146.

[16] L. Wang, C. Leckie, R. Kotagiri, and J. Bezdek, "Approximate pairwise clustering for large data sets via sampling plus extension," *Pattern Recognit.*, vol. 44, no. 2, pp. 222–235, 2011.

[17] K. Tas¸demir, Y. Moazzen, and I. Yildirim, "Geodesic based similarities for approximate spectral clustering," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Stockholm, Sweden, Aug. 2014, pp. 1360–1364.

[18] A. Strehl and J. Ghosh, "Cluster ensembles—A knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, no. 3, pp. 583–617, Mar. 2002.

[19] S. Dudoit and J. Fridlyand, "Bagging to improve the accuracy of a clustering procedure," *Bioinformatics*, vol. 19, no. 9, pp. 1090–1099, 2003.

[20] A. Topchy, A. K. Jain, and W. Punch, "A mixture model for clustering ensembles," in *Proc. SIAM Int. Conf. Data Min*., 2004, pp. 379–390.

[21] A. L. N. Fred and A. K. Jain, "Combining multiple clusterings using evidence accumulation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 835–850, Jul. 2005.

[22] F. Tung, A. Wong, and D. A. Clausi, "Enabling scalable spectral clustering for image segmentation," *Pattern Recognit.*, vol. 43, no. 12, pp. 4069–4076, 2010.

[23] J. Jia, X. Xiao, and B. Liu, "Similarity-based spectral clustering ensemble selection," in *Proc. 9th Int. Conf.*