

## Accepted Manuscript

A new cluster tendency assessment method for fuzzy co-clustering in hyperspectral image analysis

Nha Van Pham, Long The Pham, Thao Duc Nguyen,  
Long Thanh Ngo

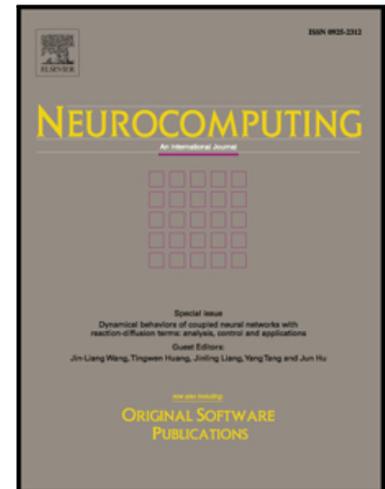
PII: S0925-2312(18)30449-1  
DOI: [10.1016/j.neucom.2018.04.022](https://doi.org/10.1016/j.neucom.2018.04.022)  
Reference: NEUCOM 19482

To appear in: *Neurocomputing*

Received date: 23 October 2017  
Revised date: 23 February 2018  
Accepted date: 13 April 2018

Please cite this article as: Nha Van Pham, Long The Pham, Thao Duc Nguyen, Long Thanh Ngo, A new cluster tendency assessment method for fuzzy co-clustering in hyperspectral image analysis, *Neurocomputing* (2018), doi: [10.1016/j.neucom.2018.04.022](https://doi.org/10.1016/j.neucom.2018.04.022)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



**Highlights**

- SACT is proposed to determine the suitable number of clusters and centroids.
- SHSI-IPD model is proposed to reduce the size of the input data of SACT.
- Experiments were conducted on the big data such as the hyperspectral images.
- The algorithms FCM, FCoC and IVFCoC were combined with SACT in experiments.
- Eight indicators are used to support cluster quality assessment of SACT algorithms.

ACCEPTED MANUSCRIPT

# A new cluster tendency assessment method for fuzzy co-clustering in hyperspectral image analysis

Nha Van Pham<sup>a,b,\*</sup>, Long The Pham<sup>a</sup>, Thao Duc Nguyen<sup>b</sup>, Long Thanh Ngo<sup>a</sup>

<sup>a</sup>*Institute of Simulation Technology, Le Quy Don Technical University,  
236 Hoang Quoc Viet, Hanoi, Vietnam*

<sup>b</sup>*MIST Institute of Science and Technology, 17 Hoang Sam, Hanoi, Vietnam*

---

## Abstract

The assessment of cluster tendency is a method determining whether a considering data-set contains meaningful clusters. The raised questions often are: How many clusters is the data-set reasonably partitioned into and How is the data-set disposed? In this paper, we proposed a new assessment method of cluster tendency which is called Silhouette-Based Assessment of Cluster Tendency (SACT). The SACT algorithm appraises the cluster tendency of the data-set in terms of the number of clusters and the initial prototypes which can be used to simultaneously determine the suitable number of clusters and the prototypes. The information of the suitable number of clusters and the prototypes helps the clustering algorithms to improve the performance. The hyperspectral image analysis is one of the complex problems which need to improve the speed of the SACT algorithm by using the Image Patch Distance technique for sparse hyperspectral image representation, i.e., reducing the size of the input data of the SACT algorithm. Experiments were conducted on some labeled synthetic data sets, color images and hyperspectral images. The proposed algorithm exhibited high performance, reliability and accuracy compared to previously algorithms in the assessment of cluster tendency.

*Keywords:* Cluster analysis, assessment of cluster tendency, hyperspectral image analysis, fuzzy co-clustering, fuzzy clustering.

---

## 1. Introduction

Clustering is an unsupervised learning method that divides a data-set into the meaningful and useful clusters. Some well-known clustering algorithms are studied and widely applied to various fields like K-Means [1], Expectation Maximization (EM) [2], Hierarchical clustering (HC) [3], Spectral clustering [4], Co-clustering [5] and fuzzy clustering techniques such as Fuzzy C-Means (FCM) [6, 49], type-2 fuzzy clustering [7], fuzzy co-clustering [8, 47],

---

\*I am corresponding author

*Email addresses:* famvannha@gmail.com (Nha Van Pham), longpt@mta.edu.vnn (Long The Pham), thaodr@gmail.com (Thao Duc Nguyen), ngotlong@mta.edu.vnn (Long Thanh Ngo)

type-2 fuzzy co-clustering [9]. These techniques have continually been improved and applied to the areas of real-life problems such as biomedicine [10], Web data analysis [11], pattern recognition [12], query information [13], data classification [14].

There are two main problems in application of clustering algorithms. The first one is that these algorithms use alternating minimization methods to solve non-convex optimization problems to find the cluster solutions [16]. These algorithms require a set of initial centroids to start searching and often terminating at the optimal solution which strongly depends on the initial centroids. Therefore, the final solution is sensitive to the initial centroids. Due to its simplicity, the random initialization method has been widely used. Moreover, the random initialization method only works stable when the randomly initial centroids are close to a good solution. Therefore, how to select the initial centroids is one of the most important problems in data clustering. Recently, the FCM algorithm-based methods [49, 50, 51] have used the PSO algorithm to initialize the centroids. The second issue is that the number of clusters  $C$  needs to be determined in advance as an input of the clustering algorithms. In a real data set,  $C$  is usually unknown. In practice, the different values of  $C$  are tried in which the cluster validation indices are used to measure the clustering results and determine the best suitable value of  $C$ . Some proposed methods [15, 17, 18, 19] used the Silhouette index and methods in [9, 47] used the index of compactness and separation to find the optimal number of clusters. These methods consist of a series of the clustering procedures with the number of clusters varying from  $C_{min}$  to  $C_{max}$ . In each iteration, the clustering results are quantified by the validity indices. After finishing the algorithm, the value of the indices is used to estimate the optimal number of clusters.

The assessment of cluster tendency is a method that determines whether a considering data-set contains meaningful clusters. Bezdek et. al. [20] have introduced a technique known as tools for visual assessment of cluster tendency (VAT). This technique can determine the optimal number of clusters in the data-set by building an ordered dissimilarity image (ODI). We can estimate the optimal number of clusters by counting the number of dark blocks along the diagonal of ODI image. The VAT algorithm seems to work well for relatively small data sets ( $n \leq 1000$ ). However, for a data set of moderate size ( $n \geq 10,000$ ), the computation time for reordering the image matrix becomes expensive. To overcome these drawbacks of the VAT algorithm, some improved VAT algorithms have been proposed. Prabhu et. al. [38] proposed Enhanced VAT by producing a binary image, which can be visually assessed for the cluster tendency. Then enhanced VAT reduces the computational complexity and performs similarities with different measures of metrics that are used in an effective visual evaluation process. Huband et. al. [21] proposed a bigVAT algorithm which used to solve the large data problem suffered by the VAT algorithms. Automated VAT algorithms [22, 23] was combined with a path-based distance transform and Spectral VAT [37] based on the difference between diagonal blocks and off-diagonal blocks which were proposed to automatically determine the number of clusters use ODI image. Havens [24] proposed iVAT by using a graph-theoretic distance transform to improve the effectiveness of the VAT algorithm which reduces the computational complexity of the algorithm iVAT from  $O(N^3)$  to  $O(N^2)$ . Essentially, the improved VAT algorithms have overcome some of the limitations of VAT and have automatically identified the number of data clusters. In

general, the improved VAT algorithms can only determine optimal clusters on small and medium data sets.

In this paper, we proposed a new assessment of cluster tendency algorithm (called SACT). The SACT algorithm is used to solve two fundamental problems in data clustering. First, we represent the data-set as a weighted graph and construct a minimum spanning tree (MST) using the Prime algorithm. Where, the vertexes correspond to the data points, the size of each edge is the Euclidean distance between two corresponding data points. Then, the SACT algorithm iteratively hash MST to  $C_{max}$  branches. In each iteration, the largest weighted edge is found in the MST to hash and removes this edge from the MST. After each hashing step, the values of the silhouette validity index are calculated and the status of branches is stored. Finish the hashing, we can rely on the values of the silhouette validity index in hashing steps to determine the optimal number of clusters and rely on the status of branches to determine the initial centroids. In order to reduce the size of the input data of the SACT algorithm, we proposed a sparse hyperspectral image representation model using Image Patch Distance Technique. Experimental results on several synthetic data sets, color images and hyperspectral images demonstrate the effectiveness of the method for determining the optimal number of clusters and obtaining better clustering results.

The organization of the rest of the paper is as follows: some related works are introduced in Section 2, the proposed method is presented in Section 3. Section 4 benchmark the proposed method against the related ones on several data sets. We finally conclude in Section 5 with discussions.

## 2. Background

### 2.1. Visual Assessment of Cluster Tendency method

Bezdek et al. [20] introduced a technique known as tools for visual assessment of cluster tendency (VAT). This study presented an idea which can determine the suitable number of data clusters by building a simple process, easy to understand using spanning tree finding techniques based on Euclidean distance measurement to build an ordered dissimilarity image (ODI). From ODI, we can use our senses to determine the number of data clusters. Original VAT is described summarized as follows.

Let  $X = \{x_i | x_i \in R^D\}, i=1, \dots, N$  be a data set which contains  $N$  objects in the  $D$ -dimensional space.  $R = \{R_{ij}\}_{N \times N}$  is a pairwise matrix of dissimilarities between objects, each element of which  $R_{ij}$  is the dissimilarity between objects  $x_i$  and  $x_j$ . In data clustering, Euclidean distance is often used to calculate the difference between data objects. In this case,  $R_{ij}$  is the squared Euclidean distance by the following formula.

$$R_{ij} = \|x_i - x_j\|^2 = \sum_{d=1}^D (x_{id} - x_{jd})^2 \quad (1)$$

By convention above, for all  $i, j$  with  $1 \leq i, j \leq N$ ,  $R$  satisfies the following conditions:

$$\begin{cases} R_{ij} \geq 0 \\ R_{ij} = R_{ji} \\ R_{ii} = 0 \end{cases} \quad (2)$$

The objective of the VAT algorithm is to rearrange the rows and columns of the matrix  $R$  according to the desired order, and then displays this matrix as a density image. Then, the VAT algorithm used visual to determine the number of clusters. Let  $K$ ,  $I$  and  $J$  be the sets of integers corresponding to the index of the data objects in the data set.  $P$  is an array of integer values corresponding to the index of the row or column of the correlation matrix  $R$  will be rearranged. The VAT algorithm is represented in Algorithm 2.1.

---

**Algorithm 1** VAT algorithm: Determine the number of clusters by density image data

---

**Input:** Correlation matrix  $R$  size  $N \times N$ .

**Output:** The number of data clusters.

**Step 1:** Set  $K = \{1, 2, \dots, N\}$ ;

Select  $(i, j) \in \arg \max\{d_{pq}\}$  with  $p, q \in K$ .

Set  $P(1) = \{i\}$ ;  $I = \{i\}$  and  $J = K - \{i\}$ .

**Step 2:** For  $t = 2, 3, \dots, N$  Do

Select  $(i, j) \in \arg \max\{d_{pq}\}$  with  $p \in I, q \in J$ ;

Set  $P(t) = j$ ; replace  $I \leftarrow I \cup \{j\}$  and  $J \leftarrow J - \{j\}$ .

**Step 3:** Rearrange the order of the rows and columns of the matrix  $R$  in the order of the indices in  $P$  matrix, we obtain a new correlation matrix  $D$  with size  $N \times N$  as follows.

$D = [d_{ij}] = [R_{P(i)P(j)}]$ ,  $1 \leq i, j \leq N$ .

**Step 4:** Display  $D$  as a density image size  $N \times N$ . Based on this image to determine the number of data clusters.

---

The VAT algorithm rearranges the pairwise distance values in a similar way to find the minimum spanning tree in the weighted graph by Prim's algorithm. The main differences between the VAT algorithm and Prim's algorithm are that: (a) The VAT algorithm does not focus on representing the MST, but only find the orders in which the vertices are added; and (b) VAT used a selection of the initial vertex that depends on the maximum edge weight in the underlying complete graph. Using the vertices of the largest weighted edge as the initial points to avoid forming the zigzagged paths.

*Example 1:* A sample data set with the image of an original density matrix, the graph of MST and the image of the arranged density matrix are shown in Fig. 1. There are five dark blocks on the main diagonal of the ODI image (Fig. 1.d). According to VAT, the suitable number of clusters in the sample data set is five.

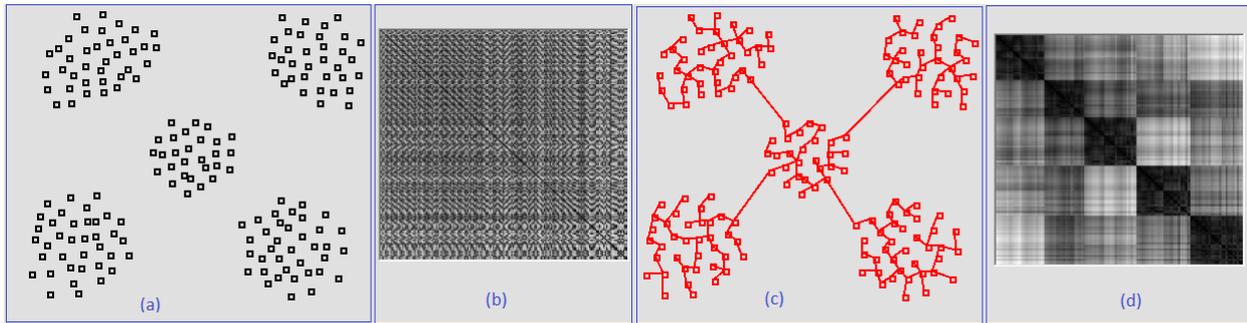


Figure 1: a) The sample data set; b) The image of an original density matrix; c) The graph of MST; d) The image of arranged density matrix  $D$

The VAT algorithm is a simple method, easy to understand, however, the VAT algorithm determines the number of clusters through visual assessment and not mention how to identify the centroids. Moreover, the larger the data is, the larger the time is to determine the minimum spanning tree and the larger the image size of ODI is, so it is difficult to observe visually. To overcome this drawback of the VAT algorithm, some improved VAT algorithms have been proposed [21], [22], [23], [24]. However, these algorithms have the high computational complexity.

## 2.2. Silhouette cluster validity index

Cluster validation allows assessing the quality of clustering results. These clustering results are usually measured in terms of compactness and separateness of clusters given by some indices. Compactness shows how closely the data are in a cluster. Separateness indicates how well-separated a cluster is from other clusters.

The Silhouette index [28] is a well-known index, which used to validate a clustering partition [17], [29]. It is based on geometrical considerations and combines ideas of both compression and separation of the clusters. We have used this index as the fitness function to select the candidate clusters.

Consider an object  $x_i$  that belongs to cluster  $C_a$ . The average dissimilarity of  $x_i$  to all of the other objects of  $C_a$  is denoted by  $a(x_i)$ . Next, let us consider cluster  $C_b$ . The average dissimilarity of  $x_i$  to all of the objects of  $C_b$  will be called  $d(x_i, C_b)$ . For all clusters  $C_b \neq C_a$ , the cluster with minimal dissimilarity to  $x_i$  is selected, i.e.,  $b(x_i) = \min_{C_b \neq C_a} d(x_i, C_b)$ . This value represents the dissimilarity of  $x_i$  to its neighbor cluster, and the silhouette  $S(x_i)$  is given by (3) as follows:

$$S(x_i) = \frac{b(x_i) - a(x_i)}{\max \{a(x_i), b(x_i)\}} \quad (3)$$

According to Equation (3), if the objects are always assigned to the closest cluster, the silhouette values will be within the interval  $[-1, +1]$ . Thus, higher silhouette values indicate a better assignment of objects to the closest clusters. Therefore, when the silhouette value is equal to zero, it is unclear to which cluster the object should be assigned, i.e., either the

current one or to a neighboring cluster. Finally, if cluster  $C_a$  is a singleton, then  $S(x_i)$  is not defined, and the most natural choice is to set  $S(x_i) = 0$ .

### 2.3. Hyperspectral Image Patch Distance Technique

The technology generating hyperspectral images combine the two latest technologies in imagery and spectroscopy, through the hyperspectral image sensor, sequential imagery data can be acquired with a high spectral resolution, which benefits the theoretical research on the hyperspectral data analysis in various fields [36]. Nowadays, this technology has widely been used in different fields, such as military, search salvage and rescue, environmental monitoring, mineral exploration and public security. The advantages of the hyperspectral image are the high spectral resolution, providing spectral characteristics and spatial information simultaneously, the large band number which contains knowledge of the target spectrum for target detection and recognition. Some recent typical studies on the hyperspectral image are for target detection of minerals[39], environmental management [40], military [41], target detection [42, 43] and the hyperspectral image classification [44, 45, 46]. As a powerful model, convolutional neural networks (CNNs) have demonstrated remarkable performance in various image representation and recognition problems. Zhang et al. [53] used CNN to recognize human genders in a video stream, Guimin et al. [54] used CNN to reconstruct high-resolution images from low-resolution ones and Yu et al. [55] proposed an efficient CNN architecture to boost its discriminative capability for hyperspectral image classification. Due to the insufficient training samples of hyperspectral image data, therefore, the unsupervised hyperspectral image clustering still have challenges to be effectively applied to the hyperspectral image analysis.

In the hyperspectral image, pixel objects within a near neighborhood are usually captured from the rectangle made by the similar materials, i.e. Their spectral characteristics are highly correlated [31]. Based on this fact, we exploit the spatial neighborhood to combine spectral and spatial-contextual information. The content of image patch distance technique (IPD) [32] is presented below.

We first consider a  $w \times w$  spatial window with central pixel  $x_{ij}$ , where  $w$  is an odd positive integer.

Let  $\Phi(x_{ij}) = \{x_{pq} | p = i - k, \dots, i, \dots, i + k; q = j - k, \dots, j, \dots, j + k\}$  is the spatial neighbors  $w^2$ -pixel set consisting of  $x_{ij}$ , in which  $k = (w - 1) / 2, i - k \leq i \leq k, j - k \leq j \leq k$ . Examples of the spatial neighbor set with size of  $3 \times 3$  are shown in Fig. 2, in which each square grid represents a pixel vector in the hyperspectral images.

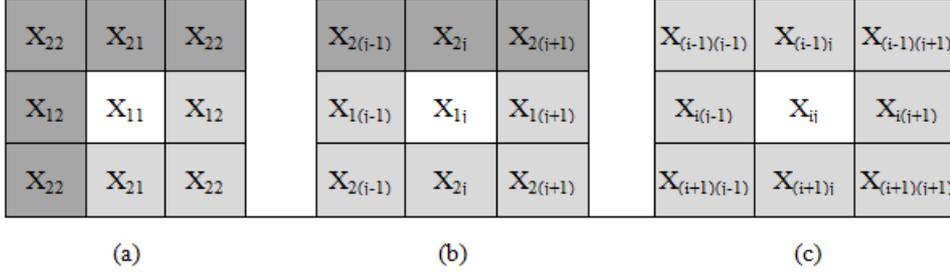


Figure 2: Examples of the spatial neighbor set of three cases with the central pixel. (a) At the corner. (b) On the edge. (c) In the image

Let  $a_l$  and  $b_l$  be the  $l^{\text{th}}$  element of the pixel sets  $\Phi(x_{ij})$  and  $\Phi(x_{pq})$ , respectively, i.e.,  $\Phi(x_{ij}) = \{a_1, a_2, \dots, a_{w^2}\}$ ,  $\Phi(x_{pq}) = \{b_1, b_2, \dots, b_{w^2}\}$ . The distance between pixel  $a_l$  and pixel set  $\Phi(x_{ij})$  is commonly defined as  $d(a_l, \Phi(x_{pq})) = \min_{b \in \Phi(x_{pq})} d(a_l, b)$  [33]. Then, we can obtain a scalar distance measure between two pixels  $a_l$  and  $b_l$  as follows,

$$d_s(a_l, b_l) = \max \left( \min_{b \in \Phi(x_{pq})} d(a_l, b), \min_{a \in \Phi(x_{ij})} d(a, b_l) \right) \quad (4)$$

where  $d(a, b)$  is a spectral similarity function comparing  $a$  to  $b$ .

With the previous definitions, we give the definition of new similarity measure between the observed pixels  $x_{ij}$  and  $x_{pq}$  as follows,

$$\begin{aligned} d_{IPD}(x_{ij}, x_{pq}) &= \sum_{l=1}^{w^2} d_s(a_l, b_l) \\ &= \sum_{l=1}^{w^2} \left\{ \max \left( \min_{b \in \Phi(x_{pq})} d(a_l, b), \min_{a \in \Phi(x_{ij})} d(a, b_l) \right) \right\} \end{aligned} \quad (5)$$

### 3. Silhouette-Based Assessment of Cluster Tendency algorithm

#### 3.1. SACT algorithm

The assessment of cluster tendency is a method that determines whether a considered data set contains meaningful clusters. In this section, we proposed a new assessment of cluster tendency method which is called Silhouette-Based Assessment of Cluster Tendency algorithm (SACT). The SACT algorithm assesses the cluster tendency in terms the number of clusters which used Silhouette index to determine the suitable number of clusters. The SACT algorithm was inspired from VAT algorithm [20]. VAT is considered as a simple method, easy to understand, however, the VAT algorithm determines the number of clusters through visual assessment. The SACT algorithm is presented as follows.

At first step, we build the weighted graph and the minimum spanning tree (MST). Then, the SACT algorithm carries out a series of MST hashing procedures with a number of branches varying from 2 to  $C_{max}$ . The data-set are modeled using the Euclidean distance to

build the weighted matrix and the weighted graph; then the Prim's algorithm is used to build the minimum spanning tree. The SACT algorithm inherited simple and can automatically determine the suitable number of data clusters and centroids simultaneously for clustering algorithms. We present a spanning tree by the adjacent list then hashing this tree by an edge which has the largest weight in the tree. The number of clusters corresponds to the number of the formed branches. At each step, Silhouette index is calculated and compared to the previous steps. If the Silhouette index in each step is higher than the previous step, then the state of the cluster distribution is saved. The result of each step is the number of clusters and cluster distribution. The status which has the largest Silhouette index corresponds to the most suitable number of clusters. The SACT algorithm is represented in Algorithm 3.1.

---

**Algorithm 2** SACT algorithm finds the suitable number of clusters and centroids

---

**Input:** Correlation matrix  $R$  with size  $N \times N$ , the maximum number of clusters  $C_{max}$ .

**Output:** The suitable number of clusters and centroids.

**Step 1.** Set  $K = \{1, 2, \dots, N\}$ ; Select  $(i, j) \in \arg \max \{d_{pq}\}$  with  $p, q \in K$ . Set  $E = \{(i, j)\}$ ;  $I = \{i\}$  and  $J = K - \{i\}$ .

**Step 2.** For  $t = 2, 3, \dots, N$

Select  $(i, j) \in \arg \max \{d_{pq}\}$  with  $p \in I, q \in J$ ; Set  $E = E \cup \{(i, j)\}$ ; replace  $I \leftarrow I \cup \{j\}$  and  $J \leftarrow J - \{j\}$ .

**Step 3.** Initialize  $T = \{E\}$ ,  $S_0 = 0$ ,  $c=1$ .

*Step 3.1.* Set  $c=c+1$ ;

*Step 3.2.* Select  $(k, i, n_i)$  with  $T=T_1, \dots, T_i, \dots, T_c$  and  $e_k \leq \arg \min\{T\}$ ,  $e_k \in T_i$ .

*Step 3.3.* Set  $T_{c+1} = \phi$ . Move  $e_{k+1}, e_{k+2}, \dots, e_{n_i} \in T_i$  to  $T_{c+1}$ .

Hash  $e_k \in T_i$ ; set  $T = T \cup T_{c+1}$ .

*Step 3.4.* Calculate Silhouette index  $S_c$ . Save the number of clusters  $C_o=c$ , Silhouette index  $S_c$  and cluster distribution status  $T$ .

*Step 3.5.* If  $C_o \geq C_{max}$  then stop and go to **Step 4**, else go back *Step 3.1*.

**Step 4.** Determine the number of clusters  $C_o$  corresponding to the maximum value of the Silhouette index  $S_c$ .

**Step 5.** Identify centroids from distributed cluster status  $T$ .

---

The SACT algorithm consists of three main phases, which are the construction of matrix  $R$  which requires  $O(N^2)$ ; build the minimum spanning tree (MST) that needs  $O(N^2)$ ; hash MST into  $C_{max}$  branches which needs  $O(C_{max}N)$ . Thus, the computational complexity of SACT is  $O(C_{max}N^2)$ , where  $N$  is the number of pixels,  $C_{max}$  is the maximum number of data clusters.

Note that after hashing, the number of clusters was increased by 1, the question here is, "The algorithm carries out hashing how many times is?". Individual data-set contains a certain number of clusters  $C_{max}$  times, so  $C_{max}$  must be determined before searching for the suitable number of clusters. The SACT algorithm will hash the tree in the range from 2 to  $C_{max}$  and the SACT algorithm find the suitable number of clusters corresponding to the highest Silhouette index.

**For the problem of determining the appropriate number of clusters:** The similarity between SACT and VAT is that both algorithms build an MST using the Prim algorithm. The difference between SACT and VAT is: VAT builds the ODI image and determines the number of clusters by visually counting the number of dark blocks on the main diagonal of ODI. Meanwhile, SACT builds MST, hash MST, and quantified the Silhouette index after each hashing. SACT determines the number of clusters by determining the time of hashing at which Silhouette index obtains the largest value.

The similarity between the SACT algorithm and the other methods [9, 15, 17, 18, 19, 47] is that these algorithms use indices to determine the optimal number of clusters. While SACT and the other methods in [15, 17, 18, 19] use the Silhouette Index, the methods in [9, 47] use the S index. The basic differences between SACT and these methods are: The previous methods must conduct a number of clustering iterations. Thus, the computational complexity of the previous methods is multiplied by the number of clustering iterations. Meanwhile, SACT builds MST and hash MST. The computational complexity of SACT is primarily focused on building the MST.

**For the problem of determining the initialization centroids:**

To identify centroids from SACT algorithm, we give the following comments. According to identifying a spanning tree, we selected the edge with the largest weight as the initialized point to avoid creating the zigzag path, then we choose the smallest value in the vicinity of the selected edges. In the case, we cannot find a smaller edge, then we had chosen a different edge which has a length greater than the other edge to take out nearby neighborhood. An order of the edges in the spanning tree is naturally formed, in which the edges of the same cluster are closely successful. Thus, clusters are localized, the question is to find where the centroid. Because the spanning tree has no cycle, so there is only one path between the two vertices of the branch. The path from the beginning to the end probably goes through intermediate points of the branches. With such reason, we have chosen the vertices of the edge in the central of the branches to generate the centroids.

Determining cluster centroids is carried out as soon as estimating the appropriate number of clusters by selecting a suitable vertex in the branches of the tree. This is considered as a simple task only on the basis of analytical reasoning about the distribution of vertices in the branches of the tree. Thus, the accuracy of the initial centroids and the computational complexity depends completely on the suitable number of clusters. Thus, the SACT algorithm can do two tasks of estimating the number of clusters and initializing centroids. Meanwhile, the previous methods in [49, 50, 51] determine the initial centroids using the particle swarm optimization algorithm.

The theoretical advantages of the SACT algorithm can be summarized as follows: 1) SACT can not only determine the number of clusters but also can determine the prototypes; 2) SACT does not produce the density image, so the memory takes less space than VAT and improved VATs. Thus, SACT is suitable with color images and hyperspectral images. 3) SACT uses compression and separation index for clustering quality. While VAT and improved VATs use image processing techniques to indirectly measure cluster quality.

To understand the SACT algorithm, we consider the following example.

*Example 1:* We need to determine the number of clusters of the data set consisting of

350 points in 2-dimensional space. The maximum number of clusters predicted by 10.

In this example, we have conducted to determine the most suitable clusters using the VAT algorithm, SACT, and the methods in [15] and [47]. We used the VAT algorithm to determine the minimum spanning tree, the original density image and the density image of VAT are shown in Fig. 3.

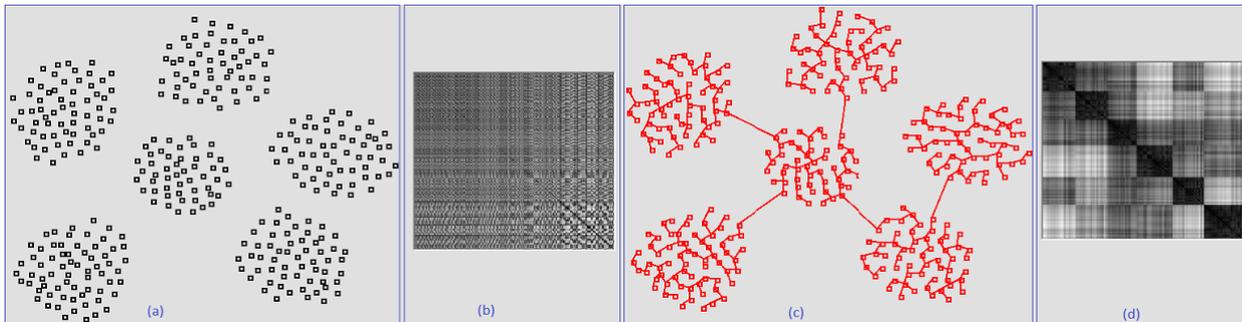


Figure 3: a) The sample data set; b) The image of an original density matrix; c) The graph of MST; d) The image of arranged density matrix D

According to the density image of VAT in Fig. 3, there are six dark blocks on the main diagonal corresponding to the number of clusters.

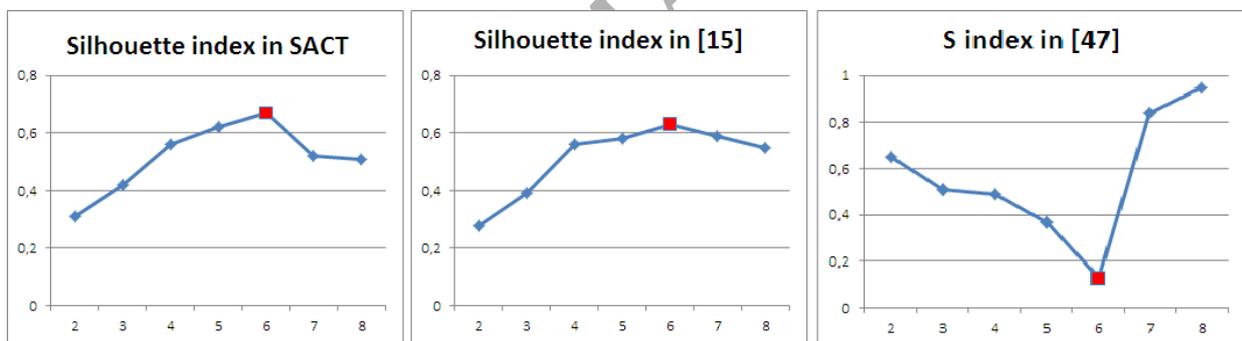


Figure 4: The diagrams of the validity indices and the number of clusters of sample data set: a) Silhouette index produced from the SACT algorithm; b) Silhouette index produced from the method in [15]; c) S index produced from the method in [47]

The results shown in Fig. 4 which produced from three methods could lead to the conclusion of the most suitable number of clusters is six clusters.

### 3.2. The sparse hyperspectral image representation model

Through the examples presented above, the SACT algorithm can be used to evaluate the number of data clusters. However, there is still a problem to have dealing with processing large datasets, such as the hyperspectral images. Because of assessing the cluster tendency of these data sets, we must construct a dissimilarity matrix based on the distance between all pixels. Thus, the size of the hyperspectral image is normally large, the size of the

dissimilarity matrix and MST will increase and the time building the dissimilarity matrix and MST will be increased considerably. For this reason, we proposed a method for reducing the size of the input of SACT by using the IPD technique.

Consider a hyperspectral image with size  $N = W \times H$ , where  $W$  and  $H$  are the width and the height of the image, respectively. Divide the image into  $M$  cells sized  $w \times w$ , where  $w$  is an odd positive integer. Assume that  $W$  and  $H$  are divisible by  $w$ , then  $M = N/w^2$ . Due to the relational properties of pixels to other pixels in their vicinity, the problem of  $N$ -pixel hyperspectral image analysis can become a problem of  $M$ -pixel hyperspectral image analysis. The dissimilarity matrix of the new problem is defined by the distances between cells using (5). The size of the new image analysis problem decreases  $w^2$  times in terms of the number of pixels. We call this method is sparse hyperspectral image representation using the image patch distance technique (SHIR-IPD). The sparse hyperspectral image model is summarized in Fig. 5.

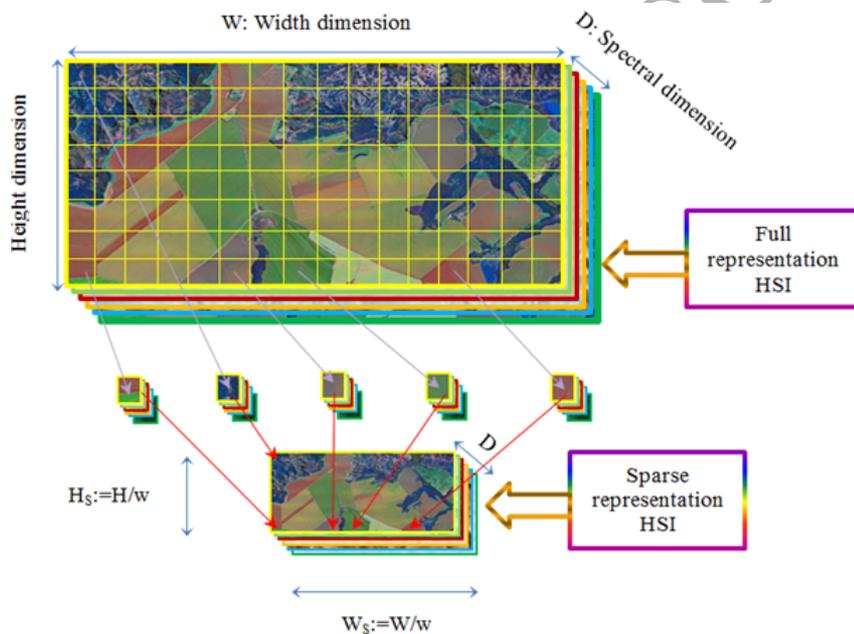


Figure 5: The sparse hyperspectral image representation model

#### 4. Results and discussion

The SACT algorithm can solve two main problems in data clustering consisting of determining the suitable number of clusters and initializing the centroids. Therefore, the performance analysis of the proposed algorithm will consist of two parts. The one is to the effectiveness in determining the suitable number of clusters. We suppose that the number of clusters is unknown in the considered data sets and use the SACT algorithm along with some previous methods [15, 20, 47] to find the candidate number of clusters. In addition, we also measure the processing time of these methods to verify performance.

Table 1: Concise information of high-dimensional synthetic data sets

Data sets	Abbreviation	No. of clusters	No. of objects	No. of features
Dim032	D1	16	1024	32
Dim064	D2	16	1024	64
Dim128	D3	16	1024	128
Dim256	D4	16	1024	256
Dim512	D5	16	1024	512
Dim1024	D6	16	1024	1024

The other is to evaluate the effectiveness of the initial cluster centroids obtained by the proposed algorithm. We use various cluster quality assessment indices to validate the effectiveness of the proposed algorithm compared with the previous methods [49, 50, 51].

#### 4.1. Performance analysis in determining the suitable number of clusters

To evaluate performance of the proposed algorithm in determining the suitable number of clusters, we conducted experiments on the high-dimensional synthetic, color image and hyperspectral image data sets. Experimental results are compared with the VAT algorithm [20] and the methods in [15] and [47]. Where, the VAT algorithm visually identifies the number of clusters. The method in [15] conducts  $C_{max}$  clustering loops. Each loop uses the FCM algorithm with number of clusters is incremented from 2 to  $C_{max}$ . The clustering result in each loop is quantified by the Silhouette index. Then, the number of clusters with the largest Silhouette value is choose. The method in [47] is similar to the method in [15] by using FCCI algorithm and the  $S$  index instead of FCM algorithm and Silhouette index, respectively. Note that the smaller  $S$  index is, the better the clustering result is, so the number of clusters corresponds to the largest value of  $S$  index.

##### 4.1.1. The labeled high-dimensional synthetic data sets

In this section, we evaluated the effectiveness of the SACT algorithm in determining the suitable number of clusters by conducting experiments on the labeled high-dimensional synthetic data sets<sup>1</sup>. These data sets have 1024 records, each of which is described by the different dimensions ( $D=32, 64, 128, 256, 512$  and 1024). Each record is labeled to one of 16 clusters. Summaries of these data sets are shown in Table 1.

We hypothesize that the number of clusters on these sets of data is unknown. Therefore, we start looking for this parameter with the hypothesis that clusters can only be maximally  $C_{max} = 20$ . We expect that the number of clusters obtained from the SACT algorithm is equal to the number of clusters obtained from the VAT algorithm, the methods proposed in [15, 47], and the number of clusters that have been labeled.

We used VAT algorithm for assessment of cluster tendency on these data sets. The results are shown in Fig. 6 including original distribution density images (ODDI) and the ordered dissimilarity image (ODI). These are gray-scale images of size  $1024 \times 1024$ .

<sup>1</sup>Speech and Image Processing Unit, School of Computing University of Eastern Finland, Clustering data sets [Online], <http://cs.joensuu.fi/sipu/datasets/>

Table 2: The execution time of methods determining the optimal number of clusters on Dim032-Dim1024

Data sets	Dim032	Dim064	Dim128	Dim256	Dim512	Dim1024
SACT	<b>1.65<sup>+</sup></b>	<b>2.01<sup>+</sup></b>	<b>2.45<sup>+</sup></b>	<b>3.34<sup>+</sup></b>	<b>5.27<sup>+</sup></b>	<b>11.26<sup>+</sup></b>
Method in [15]	7.25	10.05	14.69	19.87	28.82	41.75
Method in [47]	6.58	9.15	12.81	18.16	25.63	38.94

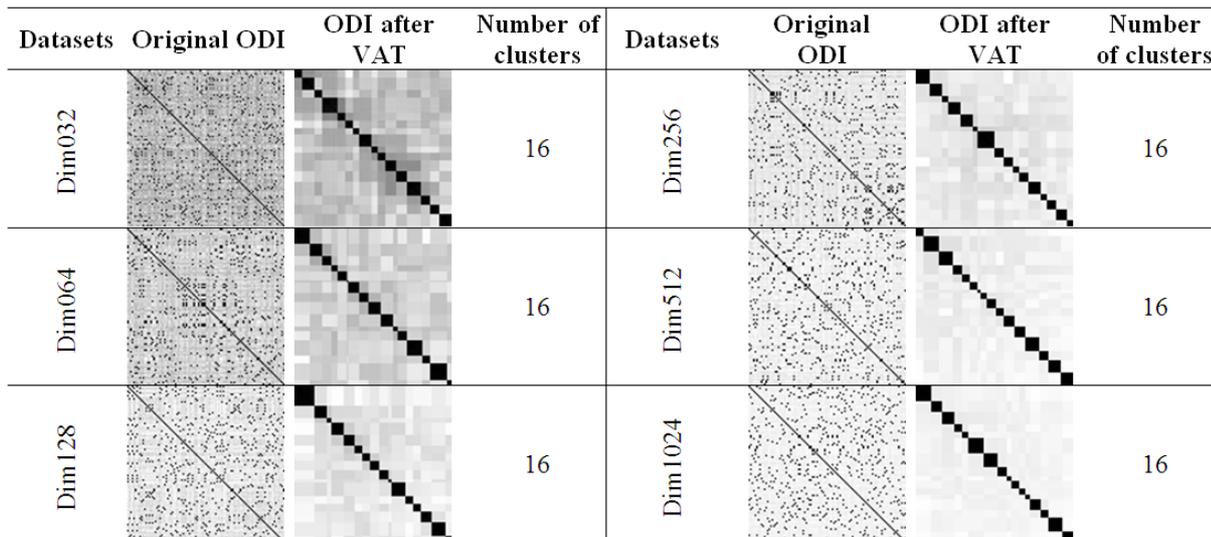


Figure 6: Experiment results to find the number of clusters using VAT algorithm

From the results in Fig. 6, we can count the number of dark blocks on the main diagonal of the ODI images. As a result, each ODI image has 16 dark blocks corresponding to the suitable number of clusters of 16.

Next, we used the SACT algorithm and the methods in [15, 47] to determine the suitable number of clusters in these data sets. The results are quantified by the Silhouette index (in SACT and [15]) and  $S$  index ([47]). The curves of indices and the number of clusters corresponding to each data set shown in Fig. 7.

On the basis of the curves in Fig. 7, the values of Silhouette index or  $S$  index reach the maximum or minimum value at the number of cluster is 16. Thus, all algorithms give the number of clusters of 16 which exactly equal to the number of the labeled classes in these data-sets.

To discuss about performance of speed in determining the number of clusters of the SACT algorithm, we measured the consumed time on the SACT algorithm and the methods in [15, 47]. Note that, in this case, we do not mention the VAT algorithm, because the VAT algorithm determines the number of clusters visually. The execution time is shown in Table 2 (in seconds). Table 2 exhibits the execution time of the SACT algorithm on the Dim032-Dim1024 data sets is lower than the methods in [15] and [47].

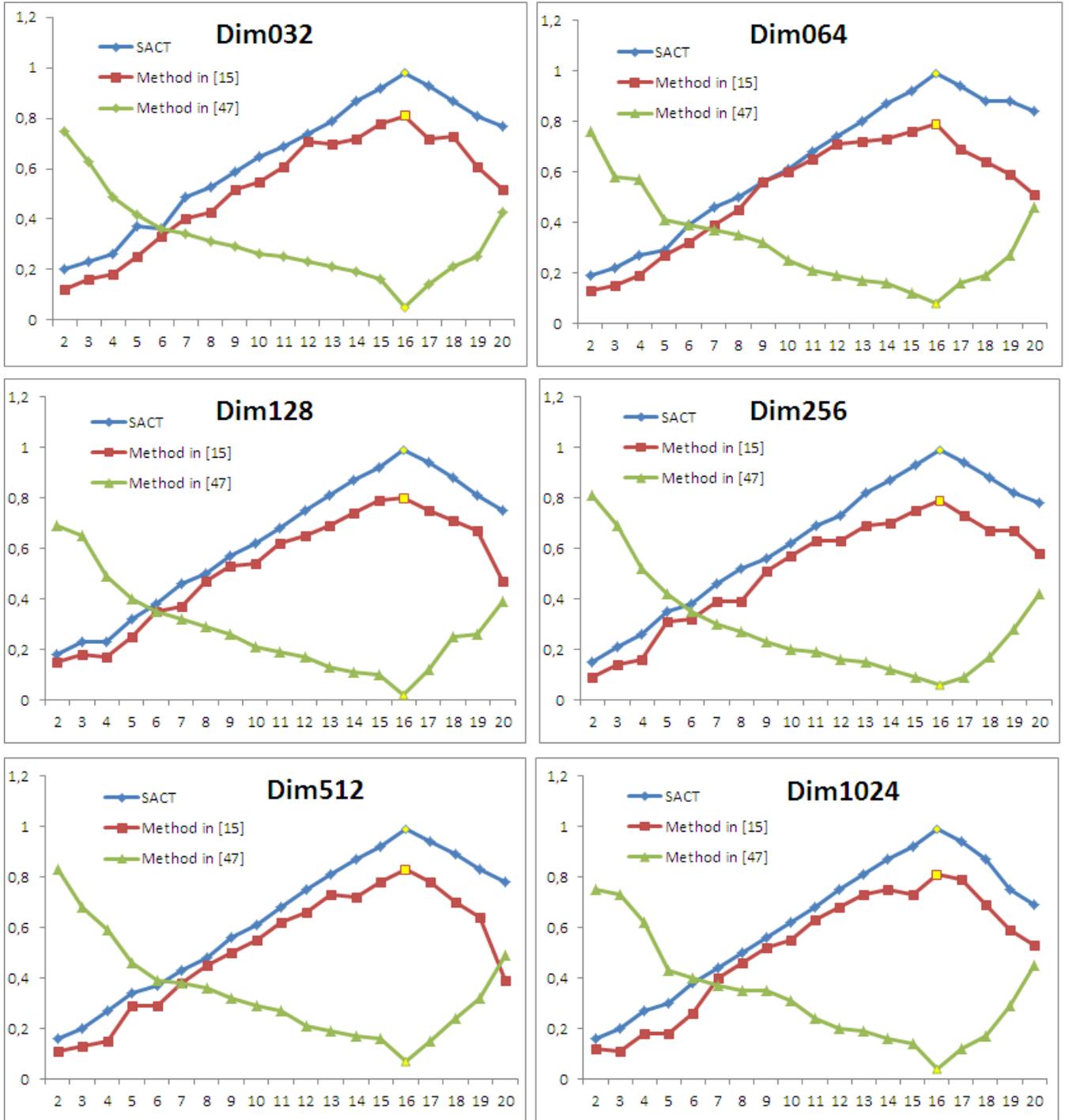


Figure 7: The curves of the indices and the number of clusters in data-sets of Dim032-Dim1024 using the SACT algorithm and the methods in [15, 47]

Table 3: The execution time of methods determining the optimal number of clusters on the color images

Data sets	14087.jpg	101055.jpg	147091.jpg	241004.jpg	42049.jpg	295087.jpg
SACT	<b>1.28<sup>+</sup></b>	<b>1.59<sup>+</sup></b>	<b>1.28<sup>+</sup></b>	<b>1.10<sup>+</sup></b>	<b>0.95<sup>+</sup></b>	<b>1.13<sup>+</sup></b>
Method in [15]	8.68	9.12	8.68	8.35	5.37	7.87
Method in [47]	7.46	7.95	7.46	7.11	4.72	6.53

Table 4: Concise information of hyperspectral image data sets

Data sets	Samples	Lines	Bands
Gulf of Mexico Wetland Sample (GoMWs)	320	660	360
Geologic Sample (Gs)	320	600	357
Agriculture & Vegetation Sample(AVs)	320	600	360

#### 4.1.2. The unlabeled data sets

Next, we used the SACT algorithm and the methods [15, 47] to assess the cluster tendency on unlabeled data sets<sup>2</sup> and hyperspectral images<sup>3</sup>. In this section, we do not present the results of the cluster tendency assessment using the VAT algorithm because the obtained result images from this method are too large ( $9600 \times 9600$  pixels for color images and  $211,200 \times 211,200$  pixels for hyperspectral images).

We randomly selected the six color images with size of  $120 \times 80$ . The results of the cluster tendency assessment of these images are represented by the curves of Silhouette index and the number of clusters as shown in Fig. 8. As the result in Fig. 8, the indices respectively take the extreme value at the number of clusters of 4, 4, 3, 3, 2 and 4.

The execution time of the SACT algorithm on the color images compared to the methods in [15] and [47] are shown in Table 3.

In the next experiments, we used three hyperspectral image data-sets. The summary of these data sets is shown in Table 4.

<sup>2</sup>The Berkeley Segmentation Dataset and Benchmark [Online], <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>

<sup>3</sup>SpecTIR's Advanced Hyperspectral & Geospatial Solutions <http://www.spectir.com/>

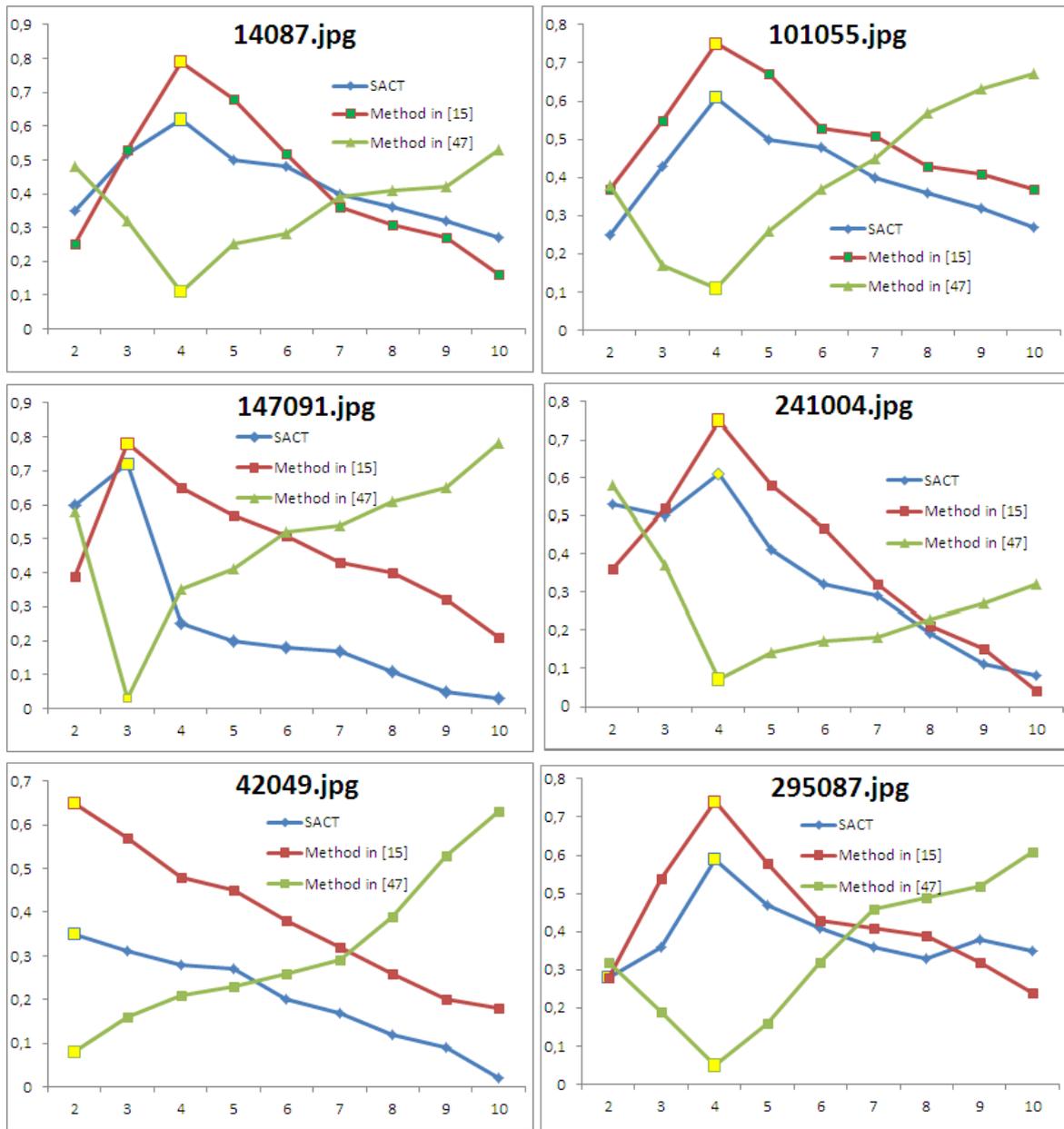


Figure 8: The characteristic curves of the indices and the number of clusters on 5 color images using SACT algorithm

Before assessing the cluster tendency, the IPD technique with different spatial windows varying of  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$ ,  $11 \times 11$ ,  $13 \times 13$  and  $15 \times 15$  is applied on the individual original hyperspectral image data-set to reduce the size of input data for the cluster tendency assessment algorithms. For example of spatial window sized  $9 \times 9$ , each pixel in the sparse data set corresponds to a cell consisting of 81 pixels in original image region. Thus, the size of the sparse image data set will be reduced 81 times compared to the original image data set. The experimental results are represented through the numbers of cluster and Silhouette

Table 5: The execution time determining the optimal number of clusters on the hyperspectral images

Data set	The use method	Using IPD technique with sizes of spatial windows varies							
		Original	3x3	5x5	7x7	9x9	11x11	13x13	15x15
GoMWs	SACT	N/A	N/A	<b>29<sup>+</sup></b>	<b>14<sup>+</sup></b>	<b>8<sup>+</sup></b>	<b>6<sup>+</sup></b>	<b>5<sup>+</sup></b>	<b>2<sup>+</sup></b>
	Method in [15]	12,688	521	157	95	58	39	30	22
	Method in [47]	10,978	480	126	65	43	32	25	18
Gs	SACT	N/A	N/A	<b>28<sup>+</sup></b>	<b>13<sup>+</sup></b>	<b>7<sup>+</sup></b>	<b>6<sup>+</sup></b>	<b>4<sup>+</sup></b>	<b>2<sup>+</sup></b>
	Method in [15]	12,239	501	153	85	54	37	29	19
	Method in [47]	10,472	443	122	81	49	32	22	16
AVs	SACT	N/A	N/A	<b>28<sup>+</sup></b>	<b>15<sup>+</sup></b>	<b>8<sup>+</sup></b>	<b>5<sup>+</sup></b>	<b>4<sup>+</sup></b>	<b>2<sup>+</sup></b>
	Method in [15]	12,496	519	155	96	58	39	26	21
	Method in [47]	10,653	468	123	83	46	32	20	17

index which are shown in Fig. 9.

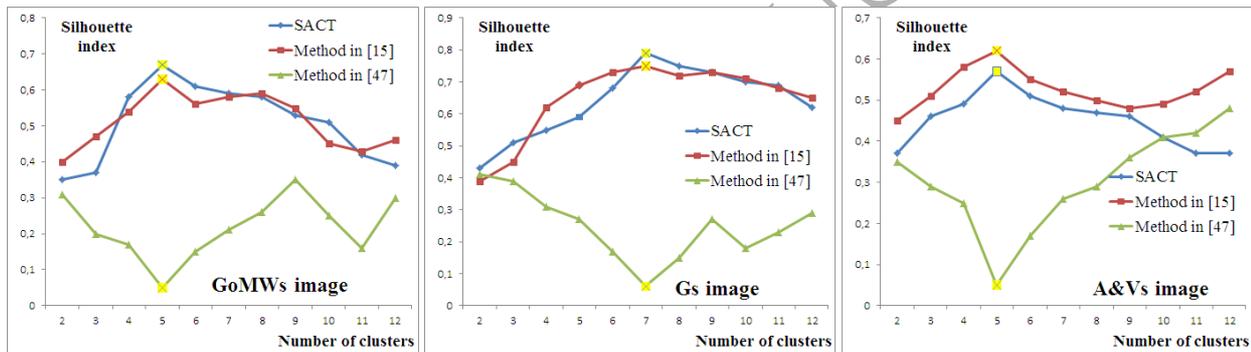


Figure 9: The characteristic curves graph of the indices and the number of clusters of hyperspectral images

The experimental results in Fig. 9 exhibits the Silhouette index takes the maximum value corresponding to the number of clusters is 7, 5 and 5 for Gs, GoMWs and AVs, respectively. Thus, the suitable number of clusters in the data-set of GoMWs, Gs, and AVs is 5, 7 and 5, respectively. The execution time of these methods is shown in Table 5.

Table 5 show that the execution time of the SACT algorithm is lower than the considered methods. However, in some cases of the large data size, the SACT algorithm cannot be performed with small spatial window due to a buffer overflow.

#### 4.1.3. Discussion about determining the suitable number of clusters

Determining the suitable number of clusters is one of the two problems of the SACT algorithm. The experimental results on the labeled data sets presented in Section 4.1.1 indicates that the number of clusters produced from the SACT algorithm is equal to the labeled number of classes and the results by running various previous methods like VAT algorithms. This demonstrates the reliability of the SACT algorithm in determining the number of clusters which has been verified through six labeled datasets.

Based on the quantitative results shown in Tables 2, 3 and 5 which were experimented on the color image dataset and the hyperspectral image datasets, the consumed time taken by the SACT algorithms is much lower than for the considered methods. That exhibits the SACT algorithm could be feasible in determining the suitable number of clusters.

#### 4.2. Performance analysis in determining the centroids

To evaluate performance of the SACT algorithm in terms of centroid determination, we have conducted experiments in pairs of FCM-SACT with FCM-PSO, FCCI-SACT with FCCI and IVFCoC-SACT with IVFCoC. The FCM-PSO algorithm [49] used particle swarm optimization (PSO) algorithm to determine the initial centroids for the FCM algorithm. Initializing centroids using the PSO algorithm is considered as a new approach applied in some recent studies [49, 50, 51, 52]. The algorithms of IVFCoC [9] and FCCI [47] are the improved fuzzy co-clustering algorithms. The terms of FCM-SACT, FCCI-SACT and IVFCoC are the combined results of the algorithms of FCM [6], IVFCoC [9] and FCCI [47] with the SACT algorithm, i.e., the algorithms of FCM, FCCI and IVFCoC used the number of clusters and centroids obtained from the SACT algorithm as input parameters.

The experimental results are represented through the validity indices of PC [25], MSE [26] and IQI [27], Recall and Precision [34], Dunn's index separation (DI), Davies-Bouldins index (DB-I), Xie and Benis index (XB-I) [35]. The clustering algorithm obtains the higher the indices of PC, IQI, Recall, Precision and D-I (called upward index) are, the lower the indices of MSE, DB-I, XB-I (called downward index) are, the better the clustering quality is.

The experimental results are presented in the table form and the color image. In particular, the table form includes the values of the indices corresponding to the clustering algorithms. The upward indices are added to the sign "+" and the downward indicators are added to the sign "-". The resulting color image consists of the pixel clusters colored by different colors.

##### 4.2.1. High-dimensional synthetic data sets

Experimental results on high-dimensional synthetic data sets are shown in Table 6.

Table 6 shows that the SACT-based clustering algorithms achieved better values of the indices than FCMPSO, FCCI and IVFCoC. In these experiments, two indices of Recall and Precision are only used in clustering the labeled datasets. Note that, these indices are closer to 1.0, the results are more similar to the ground true data.

##### 4.2.2. Color image datasets

Experimental results on color image datasets are shown in Table 7. Results in Table 7 show that the SACT-based clustering algorithms achieved better values of the indices than FCMPSO, FCCI and IVFCoC.

##### 4.2.3. Hyperspectral image datasets

Experimental results on hyperspectral image datasets are shown in Table 8 and Fig. 10.

Table 6: Experiment results on Dim032-Dim1024 using algorithms FCM-PSO, FCCI, IVFCoC and SACT

	Algorithm	PC <sup>+</sup>	MSE <sup>-</sup>	IQI <sup>+</sup>	D-I <sup>+</sup>	DB-I <sup>-</sup>	XB-I <sup>-</sup>	Prec. <sup>+</sup>	Rec. <sup>+</sup>
D1	FCM-PSO	<b>0.90</b>	11.15	0.93	<b>0.0048</b>	<b>0.82</b>	0.44	<b>0.85</b>	<b>0.86</b>
	FCM-SACT	0.89	<b>10.02</b>	<b>0.94</b>	0.0040	0.85	<b>0.36</b>	<b>0.85</b>	<b>0.86</b>
	FCCI	0.90	8.07	0.98	<b>0.0082</b>	<b>0.62</b>	0.19	0.96	<b>0.98<sup>+</sup></b>
	FCCI-SACT	<b>0.93</b>	<b>8.00</b>	<b>0.99<sup>+</sup></b>	0.0081	0.64	<b>0.17</b>	<b>0.97</b>	<b>0.98<sup>+</sup></b>
	IVFCoC	0.96	7.45	<b>0.99<sup>+</sup></b>	0.0082	0.62	0.19	0.97	<b>0.98<sup>+</sup></b>
	IVFCoC-SACT	<b>0.98<sup>+</sup></b>	<b>3.23<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0095<sup>+</sup></b>	<b>0.51<sup>+</sup></b>	<b>0.11<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>0.98<sup>+</sup></b>
D2	FCM-PSO	0.94	11.78	0.91	<b>0.0053</b>	0.95	0.52	<b>0.85</b>	<b>0.89</b>
	FCM-SACT	<b>0.95</b>	<b>11.32</b>	<b>0.92</b>	0.0046	<b>0.93</b>	<b>0.51</b>	0.83	0.88
	FCCI	0.96	5.11	<b>0.99<sup>+</sup></b>	0.0056	0.76	0.22	0.97	0.95
	FCCI-SACT	<b>0.97</b>	<b>4.32</b>	<b>0.99<sup>+</sup></b>	<b>0.0058</b>	<b>0.71</b>	<b>0.20</b>	<b>0.98<sup>+</sup></b>	<b>0.96</b>
	IVFCoC	0.99	1.19	<b>0.99<sup>+</sup></b>	0.0065	0.76	0.21	<b>0.98<sup>+</sup></b>	0.96
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>0.92<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0081<sup>+</sup></b>	<b>0.63<sup>+</sup></b>	<b>0.18<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>0.97<sup>+</sup></b>
D3	FCM-PSO	0.86	13.7	0.89	0.0035	<b>0.93</b>	<b>0.41</b>	<b>0.92</b>	0.89
	FCM-SACT	<b>0.89</b>	<b>12.1</b>	<b>0.90</b>	<b>0.0036</b>	<b>0.93</b>	0.48	0.90	<b>0.90</b>
	FCCI	0.88	24.4	0.96	0.0037	0.67	0.23	0.97	<b>0.97</b>
	FCCI-SACT	<b>0.96</b>	<b>9.25</b>	<b>0.97</b>	<b>0.0039</b>	<b>0.65</b>	<b>0.18</b>	<b>0.98</b>	<b>0.97</b>
	IVFCoC	0.98	1.19	<b>0.99<sup>+</sup></b>	0.0035	0.67	0.23	<b>0.98<sup>+</sup></b>	0.97
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>0.75<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0042<sup>+</sup></b>	<b>0.62<sup>+</sup></b>	<b>0.16<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>0.98<sup>+</sup></b>
D4	FCM-PSO	0.87	15.8	<b>0.91</b>	0.0062	<b>0.95</b>	0.37	<b>0.85</b>	0.84
	FCM-SACT	<b>0.89</b>	<b>12.2</b>	0.85	<b>0.0065</b>	0.99	<b>0.35</b>	0.83	<b>0.85</b>
	FCCI	0.90	8.22	0.97	0.0117	0.51	<b>0.09</b>	0.97	0.96
	FCCI-SACT	<b>0.95</b>	<b>6.43</b>	<b>0.98</b>	<b>0.0135</b>	0.57	0.11	<b>0.98</b>	<b>0.98</b>
	IVFCoC	0.98	0.66	<b>0.99</b>	0.0236	0.45	0.08	0.95	0.97
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>0.62<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0268<sup>+</sup></b>	<b>0.38<sup>+</sup></b>	<b>0.07<sup>+</sup></b>	<b>0.97<sup>+</sup></b>	<b>0.99<sup>+</sup></b>
D5	FCM-PSO	0.83	10.70	0.91	<b>0.0037</b>	<b>1.19</b>	<b>0.46</b>	0.92	0.92
	FCM-SACT	<b>0.90</b>	<b>9.45</b>	<b>0.92</b>	0.0033	1.27	<b>0.46</b>	<b>0.93</b>	<b>0.94</b>
	FCCI	0.93	6.42	0.98	0.0049	0.53	0.13	0.93	0.95
	FCCI-SACT	<b>0.95</b>	<b>4.32</b>	<b>0.99<sup>+</sup></b>	<b>0.0053</b>	<b>0.50</b>	<b>0.09<sup>+</sup></b>	<b>0.95</b>	<b>0.95</b>
	IVFCoC	0.91	0.84	<b>0.99<sup>+</sup></b>	0.0057	0.51	0.12	0.97	0.96
	IVFCoC-SACT	<b>0.98<sup>+</sup></b>	<b>0.53<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0078<sup>+</sup></b>	<b>0.35<sup>+</sup></b>	<b>0.10</b>	<b>0.98<sup>+</sup></b>	<b>0.98<sup>+</sup></b>
D6	FCM-PSO	0.89	16.82	0.85	0.0011	0.92	<b>0.38</b>	<b>0.87</b>	0.84
	FCM-SACT	<b>0.91</b>	<b>15.13</b>	<b>0.89</b>	<b>0.0014</b>	<b>0.89</b>	0.42	0.86	<b>0.85</b>
	FCCI	0.95	7.02	<b>0.98</b>	0.0037	0.74	0.19	0.96	0.93
	FCCI-SACT	<b>0.97</b>	<b>5.17</b>	<b>0.98</b>	<b>0.0045</b>	<b>0.67</b>	<b>0.16</b>	<b>0.98</b>	<b>0.98</b>
	IVFCoC	0.97	3.9	<b>0.99<sup>+</sup></b>	0.0067	0.73	0.19	0.98	<b>0.99<sup>+</sup></b>
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>1.89<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.0098<sup>+</sup></b>	<b>0.54<sup>+</sup></b>	<b>0.06<sup>+</sup></b>	<b>0.99<sup>+</sup></b>	<b>0.99<sup>+</sup></b>

Table 7: Experiment results on color images using algorithms FCM-PSO, FCCI, IVFCoC and SACT

Data sets	Algorithms	PC <sup>+</sup>	MSE <sup>-</sup>	IQI <sup>+</sup>	D-I <sup>+</sup>	XB-I <sup>-</sup>	DB-I <sup>-</sup>
14037	FCM-PSO	0.62	35.1	<b>0.93</b>	0.13	<b>0.62</b>	1.76
	FCM-SACT	<b>0.65</b>	<b>34.7</b>	0.92	<b>0.15</b>	<b>0.62</b>	<b>1.75</b>
	FCCI	0.98	28.8	<b>0.98<sup>+</sup></b>	0.27	1.07	1.97
	FCCI-SACT	<b>0.99<sup>+</sup></b>	<b>27.3</b>	<b>0.98<sup>+</sup></b>	<b>0.37</b>	<b>0.79</b>	<b>1.62</b>
	IVFCoC	<b>0.99<sup>+</sup></b>	26.8	<b>0.98<sup>+</sup></b>	0.45	0.38	0.95
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>25.1<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>0.92<sup>+</sup></b>	<b>0.26<sup>+</sup></b>	<b>0.87<sup>+</sup></b>
101055	FCM-PSO	0.67	<b>21.5</b>	<b>0.86</b>	0.12	0.75	1.59
	FCM-SACT	<b>0.68</b>	24.2	0.82	<b>0.19</b>	<b>0.73</b>	<b>1.55</b>
	FCCI	<b>0.98</b>	25.1	0.86	0.25	1.23	1.89
	FCCI-SACT	0.97	<b>24.7</b>	<b>0.90</b>	<b>0.32</b>	<b>1.18</b>	<b>1.75</b>
	IVFCoC	0.98	21.8	0.93	0.63	0.29	0.65
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>20.1<sup>+</sup></b>	<b>0.95<sup>+</sup></b>	<b>0.96<sup>+</sup></b>	<b>0.22<sup>+</sup></b>	<b>0.52<sup>+</sup></b>
147091	FCM-PSO	<b>0.62</b>	30.1	0.81	0.23	<b>0.85</b>	1.85
	FCM-SACT	0.58	<b>29.9</b>	<b>0.82</b>	<b>0.46</b>	0.86	<b>1.76</b>
	FCCI	0.97	24.5	0.87	0.82	1.24	1.87
	FCCI-SACT	<b>0.98</b>	<b>23.5</b>	<b>0.92</b>	<b>0.98</b>	<b>1.02</b>	<b>1.66</b>
	IVFCoC	<b>0.99<sup>+</sup></b>	22.8	0.96	0.92	0.13	0.39
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>19.2<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>1.13<sup>+</sup></b>	<b>0.11<sup>+</sup></b>	<b>0.28<sup>+</sup></b>
241004	FCM-PSO	0.66	49.2	0.73	0.10	<b>0.93</b>	<b>1.79</b>
	FCM-SACT	<b>0.67</b>	<b>45.9</b>	<b>0.74</b>	<b>0.11</b>	0.98	2.12
	FCCI	0.77	64.4	0.56	0.43	1.27	1.87
	FCCI-SACT	<b>0.88</b>	<b>52.5</b>	<b>0.91</b>	<b>0.87</b>	<b>1.01</b>	<b>1.13</b>
	IVFCoC	<b>0.99<sup>+</sup></b>	50.1	0.94	0.91	0.23	0.55
	IVFCoC-SACT	0.98	<b>36.2<sup>+</sup></b>	<b>0.97<sup>+</sup></b>	<b>1.16<sup>+</sup></b>	<b>0.19<sup>+</sup></b>	<b>0.29<sup>+</sup></b>
420049	FCM-PSO	0.61	33.8	<b>0.95</b>	0.16	0.89	1.59
	FCM-SACT	<b>0.62</b>	<b>32.7</b>	0.92	<b>0.19</b>	<b>0.88</b>	<b>1.52</b>
	FCCI	<b>0.99<sup>+</sup></b>	25.9	0.96	0.33	1.65	1.83
	FCCI-SACT	<b>0.99<sup>+</sup></b>	<b>23.1</b>	<b>0.98<sup>+</sup></b>	<b>0.87</b>	<b>1.26</b>	<b>1.72</b>
	IVFCoC	<b>0.99<sup>+</sup></b>	25.5	0.96	0.87	0.29	0.65
	IVFCoC-SACT	<b>0.99<sup>+</sup></b>	<b>20.9<sup>+</sup></b>	<b>0.98<sup>+</sup></b>	<b>1.25<sup>+</sup></b>	<b>0.28<sup>+</sup></b>	<b>0.58<sup>+</sup></b>

Table 8: The clustering results on the hyperspectral image using the clustering algorithms FCM, FCCI, IVFCoC and SACT

Data sets	Algorithm	PC <sup>+</sup>	MSE <sup>-</sup>	IQI <sup>+</sup>	D-I <sup>+</sup>	XB-I <sup>-</sup>	DB-I <sup>-</sup>
GoMWs	FCM-PSO	<b>0.58</b>	76.2	0.90	0.11	<b>0.62</b>	1.89
	FCM-SACT	<b>0.58</b>	<b>76.1</b>	<b>0.92</b>	<b>0.27</b>	0.70	<b>1.81</b>
	FCCI	0.92	76.8	0.94	0.17	1.15	2.15
	FCCI-SACT	<b>0.95</b>	<b>72.1</b>	<b>0.97<sup>+</sup></b>	<b>0.39</b>	<b>0.92</b>	<b>1.68</b>
	IVFCoC	0.95	72.9	0.95	0.17	0.95	1.78
	IVFCoC-SACT	<b>0.97<sup>+</sup></b>	<b>69.7<sup>+</sup></b>	<b>0.97<sup>+</sup></b>	<b>0.43<sup>+</sup></b>	<b>0.47<sup>+</sup></b>	<b>0.85<sup>+</sup></b>
Gs	FCM-PSO	0.51	73.2	0.86	<b>0.21</b>	1.02	1.95
	FCM-SACT	<b>0.52</b>	<b>71.3</b>	<b>0.90</b>	0.15	<b>0.98</b>	<b>1.83</b>
	FCCI	0.91	70.1	0.85	0.05	1.10	1.94
	FCCI-SACT	<b>0.93</b>	<b>69.2</b>	<b>0.91</b>	<b>0.35</b>	<b>0.84</b>	<b>1.48</b>
	IVFCoC	0.92	72.1	0.89	0.45	0.80	0.92
	IVFCoC-SACT	<b>0.98<sup>+</sup></b>	<b>65.3<sup>+</sup></b>	<b>0.96<sup>+</sup></b>	<b>0.73<sup>+</sup></b>	<b>0.75<sup>+</sup></b>	<b>0.75<sup>+</sup></b>
AVs	FCM-PSO	0.44	<b>71.9</b>	0.88	0.09	0.68	<b>1.15</b>
	FCM-SACT	<b>0.45</b>	73.9	<b>0.89</b>	<b>0.11</b>	<b>0.63</b>	1.18
	FCCI	0.87	79.2	0.91	0.09	0.93	<b>0.96</b>
	FCCI-SACT	<b>0.92</b>	<b>76.1</b>	<b>0.94</b>	<b>0.81<sup>+</sup></b>	<b>0.76</b>	<b>0.96</b>
	IVFCoC	0.91	70.2	0.93	<b>0.59</b>	0.73	0.68
	IVFCoC-SACT	<b>0.97<sup>+</sup></b>	<b>65.2<sup>+</sup></b>	<b>0.97<sup>+</sup></b>	<b>0.59</b>	<b>0.65<sup>+</sup></b>	<b>0.48<sup>+</sup></b>

Results in the Tables 6, 7 and 8 exhibit the SACT-based clustering algorithms produce the better values of the indices than FCM-PSO, FCCI and IVFCoC. In which, the IVFCoCSACT algorithm produces the best results. The clustering results obtained from the algorithms of FCM, FCCI, and IVFCoC using the centroids from the SACT algorithm are better than some other centroids initialization methods.

#### 4.3. Discussion

The SACT algorithm was inspired by the VAT algorithm which is considered a simple and effective cluster tendency assessment algorithm. However, the limitation of VAT is the visual assessment method, so it's difficult to applied to large datasets.

The SACT algorithm can be considered as an improvement of the VAT algorithm. The SACT algorithm inherited the idea of the VAT algorithm for building a minimum spanning tree. However, the SACT algorithm uses the quantitative method of Silhouette index to determine the number of data clusters instead of using the visual assessment method as the VAT algorithm. In addition, the SACT algorithm treats on the individual data object, therefore the SACT algorithm can simultaneously predict the number of clusters and the centroids.

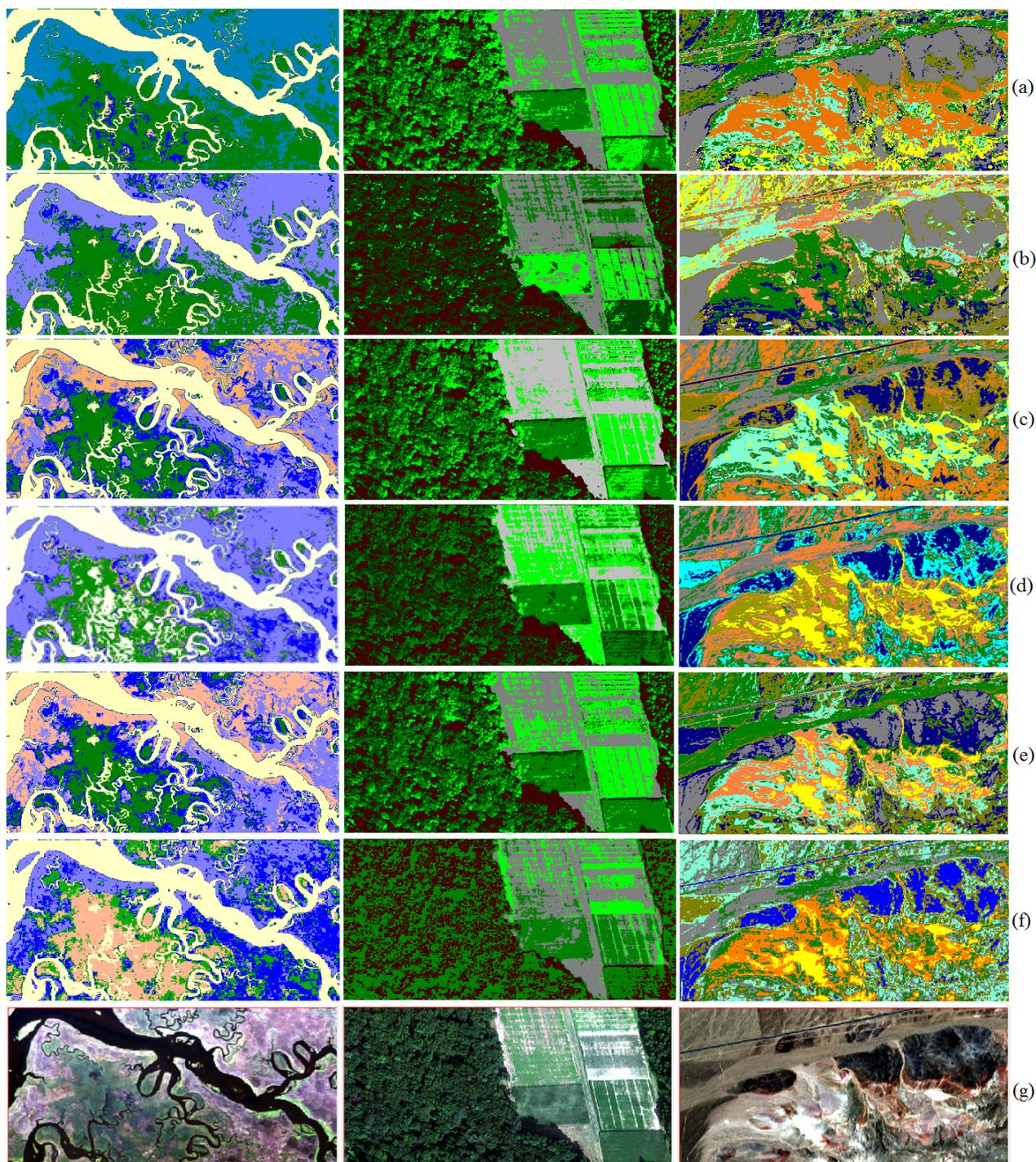


Figure 10: Hyperspectral image clustering results a) using IVFCoC-SACT; b) using IVFCoC; c) using FCCI-SACT; d) using FCCI; e) using FCM-SACT, f) using FCM-PSO; g) Three-band color composite using ENVI software

To demonstrate the correctness of the SACT algorithm, we conducted experiments on some labeled data sets. The experimental results in Fig. 6 and Fig. 7 show that the number

of clusters obtained from the SACT algorithm is exactly equal to the labeled number of classes in the datasets.

To demonstrate the reliability of the SACT algorithm in determining the centroids, we have incorporated the SACT algorithm into the clustering problem. We used the clustering algorithms of FCM, FCCI and IVFCoC with the centroids from the SACT algorithm. Then, we have to demonstrate empirical results of these algorithms are better than the clustering algorithms of FCM-PSO, FCCI and IVFCoC. To quantify the experimental results, we have used eight validity indices, including PC, MSE, IQI, Recall, Precision, D-I, DB-I, XB-I. Validity indices are the basis for determining performance of the clustering algorithms. In particular, we used the Recall and Precision indicators as the dedicated indicators for the labeled datasets. The experimental results in Tables 6, 7 and 8 show that the clustering algorithms using the centroids from the SACT algorithm produce the better results than the considered algorithms.

Although the SACT algorithm is intended to overcome some of the limitations of the VAT algorithm such as the size of input data and cluster tendency assessment. However, the computational complexity of SACT is  $O(C_{max}N^2)$ , where  $C_{max}$  is the maximum number of clusters,  $N$  is the size of input data.

Therefore, in order to be able to apply the SACT algorithm in the real-world applications such as image processing applications including color images, multi-spectral images and hyperspectral images, we have also applied a sparse hyperspectral image representation model to reduce the input data size for the SACT algorithm. The results in Table 5 exhibit somewhat the advantages of this model.

## 5. Conclusion

In this paper, we have proposed a new cluster tendency assessment method for simultaneously determining the suitable number of cluster and centroids in the hyperspectral image analysis applications which is called Silhouette-based Assessment of Cluster Tendency algorithm (SACT). The SACT algorithm assesses the cluster tendency of datasets in terms identifying the number of clusters and initializing centroids simultaneously. First, the dataset is modeled by using the Euclidean distance to build the weighted matrix and the weighted graph; then the Prim algorithm is used to build the minimum spanning tree. Next, the original minimum spanning tree was hashed into branches, each branch corresponding to one cluster. Finally, the Silhouette index is used to determine the suitable number of clusters. We also used the IPD technique to represent sparse the hyperspectral images to reduce the data input size for the SACT algorithm. In order to validate the effectiveness of SACT algorithm, a series of experiments were conducted using six labeled high-dimensional synthetic datasets, six color images and three hyperspectral image datasets.

The hyperspectral image data plays an important part in quantitative remote sensing, military, environmental management, mineral mining, biological and medical, precision agriculture applications. In the future, we will apply the SACT algorithm to conduct further applications of classification, target detection and change detection. A potential direction

is how to use the SACT algorithm together dimensionality reduction, CNNs in a framework to enhance performance of hyperspectral image classification.

## Acknowledgement

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.05-2016.09.

## References

- [1] J.B. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, pp. 281-297, 1967.
- [2] A.P. Dempster, N.M. Laird, D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society B, Vol. 39(1), pp. 1-38, 1977.
- [3] L. Rokach and O. Maimon, Clustering methods, Data mining and knowledge discovery handbook, Springer US, pp. 321-352, 2005.
- [4] J. Shi and J. Malik, Normalized Cuts and Image Segmentation, IEEE Transactions on PAMI, Vol. 22(8), pp. 888-905, 2000.
- [5] I.S. Dhillon, Co-clustering documents and words using bipartite spectral graph partitioning, Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, pp. 269-274, 2001.
- [6] J.C. Dunn, A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, Journal of Cybernetics, Vol. 3(3), pp. 32-57, 1973.
- [7] C. Hwang, F.C.H. Rhee, Uncertain fuzzy clustering: Interval type-2 fuzzy approach to C-means, IEEE Transactions on Fuzzy Systems, Vol. 15(1), pp. 107-120, 2007.
- [8] K. Kummamuru, A. Dhawale, R. Krishnapuram, Fuzzy Co-clustering of Documents and Keywords, IEEE International Conf. on Fuzzy Systems, Vol. 2, pp. 772-777, 2003.
- [9] V. N. Pham, L. T. Ngo, W. Pedrycz, Interval-valued fuzzy set approach to fuzzy co-clustering for data classification, Knowledge-Based Systems, Vol. 107, pp. 1-13, 2016.
- [10] L. Malliga, K.B. Raja, A Novel Content Based Medical Image Retrieval Technique with Aid of Modified Fuzzy C-Means Clustering (CBMIR-MFCM), Journal of Medical Imaging and Health Informatics, Vol. 6(3), pp. 700-709, 2016.
- [11] L. Yang, X.Y. Geng, H.D. Liao, A web sentiment analysis method on fuzzy clustering for mobile social media users, EURASIP Journal on Wireless Communications and Networking, pp. 1-13, 2016.
- [12] S.M.M. Golsefid, M.H.F. Zarandi, I.B. Turksen, Multi-central general type-2 fuzzy clustering approach for pattern recognition, Information Sciences, Vol. 328, pp. 172-188, 2016.
- [13] N. Pandeeswari, G. Kumar, Anomaly Detection System in Cloud Environment Using Fuzzy Clustering Based ANN, Mobile Networks and Applications, Vol. 21(3), pp. 494-505, 2016.
- [14] H. Ewbank, P. Wanke, A. Hadi-Vencheh, An unsupervised fuzzy clustering approach to the capacitated vehicle routing problem, Neural Computing and Applications, Vol. 27(4), pp. 857-867, 2016.
- [15] C. Subbalakshmi, GR. Krishna, SKM. Rao, PV. Rao, A Method to Find Optimum Number of Clusters Based on Fuzzy Silhouette on Dynamic Data Set, Procedia Computer Science, Vol. 46, pp. 346-353, 2015.
- [16] D. Xu, Y. Tian, A Comprehensive Survey of Clustering Algorithms, Annals of Data Science, Vol. 2(2), pp. 165-193, 2015.
- [17] A. Mur, R. Dormido, N. Duro, S. Dormido-Canto, J. Vega, Determination of the optimal number of clusters using a spectral clustering optimization, Expert Systems with Applications, Vol. 65, pp. 304-314, 2016.

- [18] RC. Amorim, C. Hennig, Recovering the number of clusters in data sets with noise features using feature rescaling factors, *Information Sciences*, Vol. 324, pp. 126-145, 2015.
- [19] H. Ling, J. Wu, Y. Zhou, W. Zheng, How many clusters. A robust PSO-based local density model, *Neurocomputing*, Vol. 207, pp. 264-275, 2016.
- [20] J.C. Bezdek and R.J. Hathaway, VAT: A Tool for Visual Assessment of (Cluster) Tendency, *International Joint Conference on Neural Networks*, pp. 2225-2230, 2002.
- [21] J.M. Huband, J.C. Bezdek, R. J. Hathaway, bigVAT: Visual assessment of cluster tendency for large data sets, *Pattern Recognition*, Vol. 38, pp. 1875 -1886, 2005.
- [22] L. Wang, C. Leckie, K. Ramamohanarao, J.C. Bezdek, Automatically Determining the Number of Clusters in Unlabelled Data Sets, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 21(3), pp. 335-350, 2009.
- [23] L. Wang, U.T.V. Nguyen, J.C. Bezdek, C.A. Leckie, K. Ramamohanarao, iVAT and aVAT. Enhanced Visual Analysis for Cluster Tendency Assessment, *Advances in Knowledge Discovery and Data Mining*, Vol. 6118, pp. 16-27, 2009.
- [24] T.C. Havens, J.C. Bezdek, An Efficient Formulation of the Improved Visual Assessment of Cluster Tendency (iVAT) Algorithm, *IEEE Transactions on Knowledge and Data Engineering*, pp. 813-822, 2012.
- [25] J.C. Bezdek, Cluster validity with fuzzy sets, *Journal Cybernetics*, Vol. 3, pp. 58-73, 1974.
- [26] Z. Wang, A.C. Bovik, Mean squared error: love it or leave it? A new look at signal fidelity measures, *IEEE signal processing magazine*, pp. 98-117, 2009.
- [27] Z. Wang, A.C. Bovik, A universal image quality index, *IEEE signal processing letters*, Vol. 9(3), pp. 81-84, 2002.
- [28] P.J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, Vol. 20, pp. 53-65, 1987.
- [29] C. Subbalakshmi, G.R. Krishna, S.K.M. Rao, P.V. Rao, A Method to Find Optimum Number of Clusters Based on Fuzzy Silhouette on Dynamic Data Set, *Procedia Computer Science*, Vol. 46, pp. 346-353, 2015.
- [30] J.C. Bezdek and R.J. Hathaway, VAT: A Tool for Visual Assessment of (Cluster) Tendency, *International Joint Conference on Neural Networks*, pp. 2225-2230, 2002.
- [31] Y. Chen, N. M. Nasrabadi, T. D. Tran, Hyperspectral image classification via kernel sparse representation, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 51(1), pp. 217231, 2013.
- [32] H. Pu, Z. Chen, B. Wang, G-M. Jiang, A Novel SpatialSpectral Similarity Measure for Dimensionality Reduction and Classification of Hyperspectral Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52(11), pp. 7008 - 7022, 2014.
- [33] D. P. Huttenlocher, G. A. Klanderman, W. J. Rucklidge, Comparing images using the Hausdorff distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15(9), pp. 850863, 1993.
- [34] I.S. Dhillon, S. Mallela, D.S. Modha, Information-theoretic co-clustering, *ACM IC KDDM*, pp. 8998, 2003.
- [35] W. Wang, Y. Zhang, Y., On fuzzy cluster validity indices. *Fuzzy Sets Systems*, Vol. 158, pp. 20952117, 2007.
- [36] S. Deng, Y. Xu, Y. He, J. Yin, Z. Wu, A hyperspectral image classification framework and its application, *Information Sciences*, Vol. 299, pp. 379393, 2015.
- [37] C. Leckie, L. Wang, X. Geng, J. Bezdek, K. Ramamohanarao, Enhanced Visual Analysis for Cluster Tendency Assessment and Data Partitioning, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 22(10), pp. 1401-1414, 2010.
- [38] P. Prabhu, K. Duraiswamy, Enhanced VAT for cluster quality assessment in unlabelled data sets, *Journal of Circuits, Systems and Computers*, Vol. 21(1), 1250001 (19 pages), 2012.
- [39] S. Dehnavi, Y. Maghsoudi, M. Valadanzoej, Using spectrum differentiation and combination

- for target detection of minerals, *International Journal of Applied Earth Observation and Geoinformation*, Vol. 55, pp. 9-20, 2017.
- [40] H. M. Klerk, J. Gilbertson, M. Lck-Vogel, J. Kemp, Z. Munch, Using remote sensing in support of environmental management: A framework for selecting products, algorithms and methods, *Journal of Environmental Management*, Vol. 182, pp. 564-573, 2016.
- [41] L. Zhang, Y. Zhang, Airport Detection and Aircraft Recognition Based on Two-Layer Saliency Model in High Spatial Resolution Remote-Sensing Images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 10(4), pp. 1511 - 1524, 2017.
- [42] M. Axelsson, O. Friman, T. V. Haavardsholm, I. Renhorn, Target detection in hyperspectral imagery using forward modeling and in-scene information, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 119, pp. 124-134, 2016.
- [43] G. Wang, Y. Zhang, B. He, K. T. Chong, A Framework of Target Detection in Hyperspectral Imagery Based on Blind Source Extraction, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 9(2), pp. 835 - 844, 2016.
- [44] K. Nogueira, O. A. B. Penatti, J. A. Santos, Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification, *Pattern Recognition*, Vol. 61, pp. 539-556, 2017.
- [45] V. Risojevi, Z. Babi, Unsupervised Quaternion Feature Learning for Remote Sensing Image Classification, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 9(4), pp. 1521 - 1531, 2016.
- [46] Y. Zhou, J. Lian, M. Han, Remote Sensing Image Transfer Classification Based on Weighted Extreme Learning Machine, *IEEE Geoscience and Remote Sensing Letters*, Vol. 13(10), pp. 1405 - 1409, 2016.
- [47] M. Hanmandlua, O.P. Verma, S.S., V.K. Madasu, Color segmentation by fuzzy co-clustering of chrominance color features, *Neurocomputing*, Vol. 120, pp. 235-249, 2013.
- [48] F. Saki, N. Kehtarnavaz, Online Frame-Based Clustering with Unknown Number of Clusters, *Pattern Recognition*, Vol. 57, pp. 70-83, 2016.
- [49] TM. Silva, BA. Pimentel, RMCR. Souza, ALL. Oliveira, Hybrid methods for fuzzy clustering based on fuzzy c-means and improved particle swarm optimization, *Expert Systems with Applications*, Vol. 42(17-18), pp. 6315-6328, 2015.
- [50] R. Gupta, SK. Muttoo, SK. Pal, Fuzzy C-Means Clustering and Particle Swarm Optimization based scheme for Common Service Center location allocation, *Applied Intelligence*, Vol. 47(3), pp. 624-643, 2017.
- [51] ZH. Wu, ZC. Wu, J. Zhang, An improved FCM algorithm with adaptive weights based on SA-PSO, *Neural Computing and Applications*, Vol. 28(10), pp. 3113-3118, 2017.
- [52] S. Milner, C. Davis, H. Zhang, J. Llorca, Nature-inspired self-organization, control, and optimization in heterogeneous wireless networks, *IEEE Transactions on Mobile Computing*, Vol. 11(7), pp. 1207-1222, 2012.
- [53] H. Zhang, X. Cao, J.K.L. Ho, T.W. S. Chow, Object-Level Video Advertising: An Optimization Framework, *IEEE Transactions on Industrial Informatics*, Vol. 13(2), pp. 520-531, 2017.
- [54] L. Guimin, W. Qingxiang, Q. Lida, H. Xixian, Image super-resolution using a dilated convolutional neural network, *Neurocomputing*, Vol. 275, pp. 1219-1230, 2018.
- [55] S. Yu, S. Jia, C. Xu, Convolutional neural networks for hyperspectral image classification, *Neurocomputing*, Vol. 219, pp. 88-98, 2017.



**Nha Van Pham** received the B.Sc degree and the M.Sc degree in Computer Science from Le Quy Don University, Hanoi, Vietnam, in 1999 and 2009. He is currently working toward the PhD degree in the Institute of Science and Technology MIST, Hanoi, Vietnam. His research interests include pattern recognition, image processing, and data mining.



**Long The Pham** received the PhD and Ph.DSc degrees from Belorussian State University, Minsk, in 1982 and 1987, respectively. He is currently a Professor with the Department of Information Technology, Le Quy Don Technical University, Hanoi, Vietnam. His current research interests include optimization, fuzzy logic and virtual reality. He is also Vice President of Vietnam Mathematics Society.



**Thao Duc Nguyen** received the B.Sc, M.Sc and PhD degree in Information technology from Russian State University for the Humanities, Moscow, Russia, in 1996 and 2002. He is currently working as expert of Information systems in the Department of Advisory and Planning, Institute of Science and Technology, Ministry of National Defence Vietnam, Hanoi, Vietnam. His current research interests include AI, VR, Big data and Information system new generation.



**Long Thanh Ngo** received the M.Sc, Ph.D degrees from Le Quy Don Technical University, Vietnam, in 2003 and 2009, respectively. He is currently a Associate Professor with the Department of Information Technology and Institute of Simulation Technology, Le Quy Don Technical University, Hanoi, Vietnam. His current research interests include computational intelligence, type-2 fuzzy logic, pattern recognition and image processing.

ACCEPTED MANUSCRIPT