Journal of Information Hiding and Multimedia Signal Processing ©2020 ISSN 2073-4212
Ubiquitous International Volume **11**, Number **4**, December **2020**

A Fuzzy Approach of Large Size Remote Sensing Image Clustering

Trung Nguyen Tu*
Department of Information Technology, ThuyLoi University, 175 Tay Son, Hanoi, Vietnam

Received August 2020; revised October 2020

Abstract. *Remote sensing image segmentation is a very important stage in remote sensing image processing. In many different segmentation techniques such as KMeans, C-Means, Watersed .... KMeans is one of the widely used algorithms for remote sensing image segmentation. However, This algorithm considers only pixels that belong to the nearest cluster. The Fuzzy C-Means algorithm fixed the problem of KMeans algorithm by considering that a pixel can belong to multiple clusters with corresponding dependency. However, the Fuzzy C-Means algorithm has a very slow execution speed. In particular, the Fuzzy C-Means algorithm has memory problems when doing clustering of large images such as remote sensing images. This article presents the new remote sensing image clustering algorithm MapReduce_Fuzzy to overcome the disadvantages of Fuzzy C-Means' computation time and memory problem when executing on large image without reducing cluster quality.*

Keywords: Remote Sensing images, Image Clustering, KMeans, Fuzzy C-Means, map_Fuzzy, reduce_Fuzzy, MapReduce_Fuzzy.

1. **Introduction.** Image segmentation (or clustering) or remote sensing image segmentation has been studied for a long time and is a subject of concern. Remote sensing images are increasingly complex in terms of size, number of spectral channels and the level of the detail level of the image. There are many different segmentation methods like KMeans, morphology, Markov model, etc. Most methods only use the intensity of each pixel to segmentation. In [1], Balaji and colleagues presented a new method of segmenting images based on color characteristics from images with the conversion of pixels from RGB space to L*a*b* space and clustering on this space. In [5], the authors also combined fuzzy clustering algorithms and other gray level adjustment expressions to enhance the contrast of medical images. In [12], the authors used Wavelet to reduce noises for medical images. Currently, some algorithms use more contextual information in the process to reduce the complexity of segments [9]. In [14], the authors used a local approach based on Fuzzy C-Means clustering algorithm to enhance the contrast of remote sensing images. In the algorithms of KMeans family, the algorithm KMeans combines advantages: faster speed, cluster number controlling and effective clustering even with large images. Perhaps, these are the reasons why KMeans has been used widely in research and installed in remote sensing image processing softwares. However, when partitioning large remote sensing images, the convergence speed of the algorithm is still very slow. In [2], the authors proposed the algorithm CCEA to speed up the fuzzy KMeans algorithm. However, according to [7], KMeans loses the contextual characteristics (neighboring information) of each pixel when only the intensity feature is considered. Therefore, the authors proposed the 2D-KMeans algorithm with the addition of median values such as spatial parameters (local context

information) to increase clustering efficiency [7]. However, this improvement doubles the data. This reduces the speed of data processing in general ... and the speed of clustering compared to the original KMeans in particular. Disadvantages of KMeans algorithm consider only pixels belonging to the closest cluster. Fuzzy C-Means [8] algorithm fixes the problem of KMeans algorithm. However, the Fuzzy C-Means algorithm has a very slow execution speed. In particular, the Fuzzy C-Means algorithm has problems related to memory when performing clustering of large images. The increasing size and complexity of images in general and remote sensing images in particular will be a challenge for traditional data processing methods. It will be more effective if the big data processing methods are applied. Currently, with the development of information technology, the Industrial Revolution 4.0 has led to the explosion of data (Big Data). Big data and its analysis play an important role in the Information Technology world with applications of Cloud Technology, Data Mining, Hadoop and MapReduce [10]. Traditional technologies only apply to structured data while big data includes both structured, semi-structured and unstructured data. Finding he method to effectively handle big data has become big challenges in the new age and theres a great need for new processing methods. MapReduce is a highly efficient distributed data processing model that has been widely used in large data processing [4]. In[15], the authors proposed the High-dimensional Data Clustering Using K-means Subspace Feature Selection. In [16], the authors presented An Uneven Clustering Routing Protocol based on Improved K-means Algorithm for Wireless Sensor Network in Coal-mine. This paper presents the new clustering algorithm MapReduce_Fuzzy with using MapReduce model to overcome the disadvantages of Fuzzy C-Means about calculating time and the memory problem when performing clustering of large images without reducing cluster quality. In addition, the article also presents a formal representation of image clustering solution with MapReduce_Fuzzy.

2. **Related Work.**

2.1. **Overview of Remote sensing.** According to [3], remote sensing is is a science which remotely gathers information on the Earth surface. It includes sensing and recording energy released, processing, analyzing data and applying the information after analysis. Besides, most of receiving systems and remote sensing images processing follow a seven-step procedure as shown in figure 1.

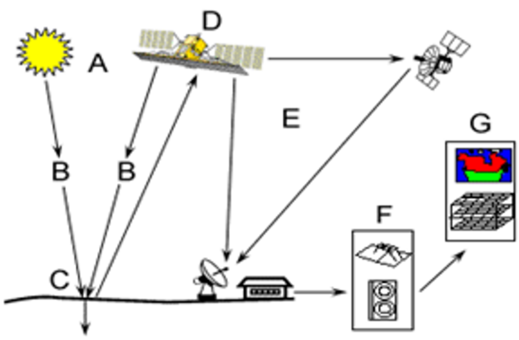

FIGURE 1. Process of gathering and processing remote sensing images [3].

In figure 1, A is energy source or bright source, B is radiance and atmosphere, C is interactive with destination object, D is energy gathered by sensor, E is energy transmission, reception and processing, F is interpretation and analysis, G is application. Remote sensing images have features: image channel, space resolution, spectrum resolution, radiant resolution and time resolution. There are many different types of remote sensing images/satellites like Landsat, SPOT, MOS, IRS, IKONOS, WORLD VIEW 2, COSMOS [11]...

2.2. **Overview of MapReduce model.** MapReduce is a model of parallel and distributed computing model that is proposed by google (Figure 2). It includes two basic functions: Map and Reduce which are defined by the user [4]. Through the MapReduce library, the program fragments the input data file. Machines include: master and worker. The master machine coordinates the operation of the MapReduce implementation process on the worker machines, the worker machines perform the Map and Reduce tasks with the data it receives. Data is structured in the form of key and value.

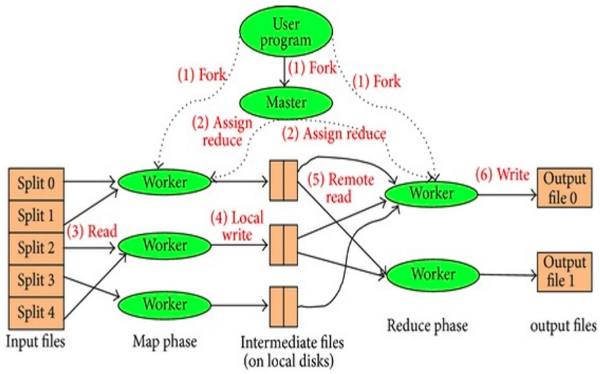

FIGURE 2. Flowchart of MapReduce model [4].

The formal representation of MapReduce model: According to [6] [12], we have the formal representation of the MapReduce model as follows:

- map: (K1 k1, V1 v1) → list(K2 k2, V2 v2)
- reduce: (K2 k2, list(V2 v2)) → list(K3 k3, V3 v3)

Where:

- K1, V1 are the input key and value types of the map function; k1, v1 are the corresponding objects with the types K1, V1.
- K2, V2 are the output key and value types of map function and still are the input key and value types of reduce function; k2, v2 are the the corresponding objects with the types K2, V2.
- K3, V3 are the output key and value types of the reduce function; k3, v3 are the the corresponding objects with the types K3, V3.

In other words, we can see:

- If k1, v1, k2, v2 are identified, we have the input and output of map function. Commonly, with text data, k1 is offset value of a data row, v1 is the content of a data row.
- If k2, v2, k3, v3 are identified, we have the input, and output of reduce function.

The formal Representation may be rewritten only with k1, v1, k2, v2, k3, v3 as follows:

$$map : (k1, v1) \to list(k2, v2) \quad (1)$$

$$reduce : (k2, list(v2)) \to list(k3, v3) \quad (2)$$

Figure 3 illustrates the diagram of the MapReduce job execution and data conversion from types (K1, V1) to types (K2, V2) and types (K2, V2) to types (K2, V3).

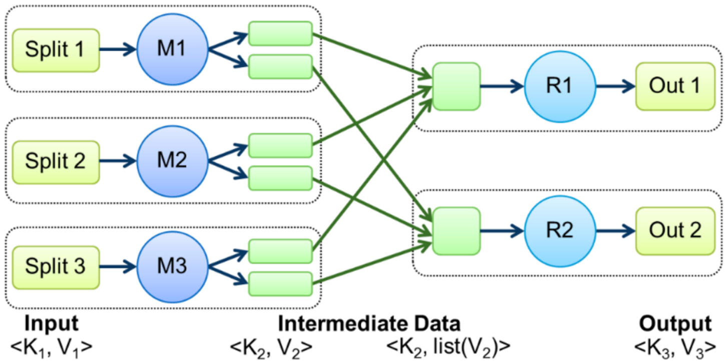

FIGURE 3. Flowchart of MapReduce model [4].

2.3. **The algorithm Fuzzy C-Means.** Fuzzy c-Means clustering algorithm [8] of fuzzy segmentation is widely used. While considering fuzzy logic set, the algorithm is developed based on k-Means clustering algorithm. In this algorithm, each pixel does not only belong to any cluster and it is represented by multiple membership of each cluster. Clustering algorithm is performed with iterating optimization of minimizing fuzzy objective function (J_m) which is defined as equation 6.

$$J_m = \sum_{i=1}^{c}\sum_{k=1}^{n} (\mu_{ik})^m d^2(x_k, V_i) \tag{3}$$

Where:

- c: number of clusters.
- n: number of pixels of image.
- μ_{ik}: membership value of i^{th} cluster of k^{th} pixel
- m: fuzzy parameter
- x_k: vector of k^{th} pixel
- V_i: center vector of i^{th} cluster
- $d^2(x_k, V_i)$: Euclidean distance between x_k and V_i

Membership (μ_{ik}) is estimated with distance between x_k and V_i and bounded as following:

$$\begin{cases} 0 \le \mu_{ik} \le 1 \\ \sum_{k=1}^{n} \mu_{ik} = 1 \\ 0 \le \sum_{k=1}^{n} \mu_{ik} \le 1 \end{cases} \tag{4}$$

Center of cluster V_i and the member matrix μ_{ik} (of the member matrix μ) can be computed by the formula:

$$V_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x_k}{\sum_{k=1}^{n} (\mu_{ik})^m} \tag{5}$$

$$\mu_{ik} \left[\sum_{j-1}^{c} \left(\frac{d(x_k, V_i)}{d(x_k, V_j)} \right)^{\frac{2}{m-1}} \right]^{-1} \tag{6}$$

Therefore, J_m can be minimized by iterating through equations (5) and (6). The first step of iterating is initializing fixed cluster number c, fuzzy parameter m, convergence threshold ε, then computing μ_{ik} and V_i using equations (5) and (6). Iterating is finished when the change of V_i between two iterations is smaller than ε. Finally, each is classified into a combination of memberships of clusters.

3. **Propose the algorithm MapReduce_Fuzzy.**

3.1. **Disadvantages of the algorithm Fuzzy C-Means.** Although the algorithm Fuzzy C-Means clusters images very efficient, however, it still has two limitations: The execution speed is very slow and getting memory problems when performing clustering of large images.

3.1.1. *The limit 1: The execution speed is very slow.* In each loop, for each data point, if the algorithm KMeans finds only the nearest cluster center, the algorithm Fuzzy C-Means must perform calculating the degrees of belonging to all cluster centers. In addition, instead of simply calculating the cluster center as KMeans, with Fuzzy C-Means, the cluster centers are calculated with the formula (5) with the participation of all points and degrees of belonging to the clusters. This leads to a very slow convergence rate of the algorithm.

3.1.2. *The limit 2: Getting memory problems when performing clustering of large images.* The algorithm Fuzzy C-Means [8] raises memory problems when executing with very large images, in particular the high spatial resolution remote sensing image. The problem arises from the membership matrix μ. According to equation (6), the size of μ is calculated as follows:

$$Size_{\mu} = c.n.8 \tag{7}$$

Where, c is the number of clusters, n is the number of pixels (size) of the image. Suppose the input image size is 2048 x 2048, the number of clusters c is 20. Then, $Size_{\mu}$ is 2048 x 2048 x 20 x 8 (Byte) = 4 x 20 x 8 (MB) = 640 (MB). The membership matrix is stored in RAM. Thus, only 1GB RAM is required to save the membership matrix in this case. However, if we want to split it into 40 clusters, then $Size_{\mu}$ is 1280 (MB)¿ 1024 (MB) = 1GB. This means that only 1GB of RAM is not enough to contain the elements of the membership matrix. And for the FCM algorithm to work, we must increase the RAM. If the image is 16000 x 16000, the number of clusters $c = 20$, $Size_{\mu}$ is 16000 x 16000 x 20 x 8 (Byte) = 39062.5 (MB) $\approx$ 39 (GB). It is easy to see, with the size and number of clusters above, even the current largest RAM for personal computers cannot have enough capacity to store this membership matrix, leading to FCM unable to execute if the matrix is Dependencies are stored in RAM. We can think of using a hard disk to store this matrix instead of using RAM. However, even with normal color images, the execution time of FCM is very slow because it takes time for data input and output operations with the hard disk. With remote sensing imagery, this time can be up to date units. This is ineffective. All this explains why FCM [8] has problems with very large size images, namely remote sensing images.

3.2. **The algorithm MapReduce_Fuzzy.** To overcome the limitations mentioned in Section 3.1, in this subsection, we propose the algorithm MapReduce_Fuzzy for clustering image data. Figure 4 is the diagram of the clustering algorithm MapReduce_Fuzzy.

According to the algorithm diagram, first, the input image is used to generate the median image (through the median filter). The data, is then converted for MapReduce processing. Next, the initiating cluster centers will be generated from this data. Next, the system will divide the data into pieces, which are processed in parallel by the MapTask (performing the calculating membership of data points to the clusters) to obtain intermediate data. After all the pieces have been processed by MapTasks, the intermediate data will be sorted, mixed, and grouped in clusters. The clustered data and membership values will then be processed by ReduceTasks to recalculate cluster centers. The system checks the convergence of cluster centers. If these centers are not converged, then system will continue to perform MapTasks and ReduceTasks. If they cluster are converged, clustered image will be performed and post-clustering performed.

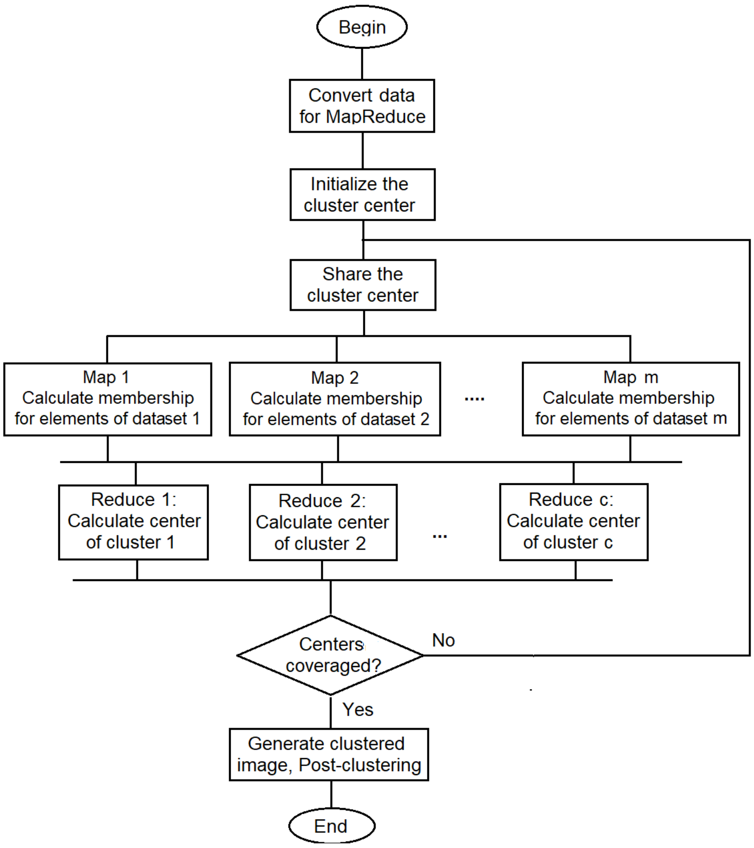

FIGURE 4. Flowchart of MapReduce model [4].

3.2.1. *Converting data for MapReduce processing.* The image data converted (include 2 components x_{ij}) into the list of rows. Each row includes: position information (row and column indices) and the list of values as vector elements representing a pixel. The reason for the position information is restoring clustered images and performing post-clustering later... Thus, the output of the clustering, data elements must include intensity information, median and corresponding positions.

3.2.2. *Formal representation of procedures map_Fuzzy and reduce_Fuzzy.* Input: Each data element x_{ij} is a set of: row and column indices, intensity components and median components: (i,j, x_{ij}). Output: The result after convergence is a set of: cluster index c and the list of elements belonging to cluster c: (i,j, x_{ij}^c). Then, the pairs of k1, v1 and k3, v3 are determined as follows:

- k1 is offset value, v1 is the content of data row (corresponding to the object x_{ij}), it means (i,j, x_{ij})
- k3 is the information of new clusters after recalculation c_{New}, v3 is the list of sets (i,j, x_{ij}) of the elements that belong to the cluster stored in k3

The Map function performs the assignment of data to the nearest cluster so k2, v2 deducing as follows:

- k2 that is the cluster index center_ind closest to x_{ij}, v2 is set (i,j, x_{ij})

At this time, the Map and Reduce procedures are be represented formaly as follows:

$$map\_Fuzzy : (offset, x_{ij}) \rightarrow list(center_ind, (i, j, x_{ij}, \mu_{ij}^c)) \tag{8}$$

$$reduce\_Fuzzy : (center_ind, list((i, j, x_{ij}, \mu_{ij}^c))) \rightarrow list(c_{New}, list(x_{ij}^{c_{New}})) \tag{9}$$

3.2.3. *The algorithm of procedures map_Fuzzy and reduce_Fuzzy.* The purpose of the map_Fuzzy algorithm is to calculate membership of the input object to each center (in the shared center set). This algorithm is presented as follows: Input: The shared center set lstCenter, key k1 is offset, value v1 is object information x_{ij}: info(x_{ij}), it means (i,j, x_{ij}). Output: lstK2V2 l list of pairs (k2,v2): k2 is cluster index, v2 is set info(i,j,x_{ij},μ_{ij}^{c}) The algorithm includes steps as follows:

- Step 1: Extract components of intensity: x_{ij}
- Step 2: Browse cen_ind = 0 to lstCenter.length
  - Step 2.1: Calculate μ_{ij}^{c}
  - Step 2.2: Initialize (k2,v2)
    - Step 2.2.1: Assign k2 = cen_ind
    - Step 2.2.2: Assign v2 = v1
  - Step 2.3: Add (k2,v2) to lstK2V2
- Step 3: return lstK2V2

The purpose of reduce_Fuzzy algorithm is to recalculate the new cluster center value from the data points and corresponding membership values. This algorithm is presented as follows: Input: key is index of cluster cen_ind, value is the list of objects x_{ij} corresponding memberships to cluster with index cen_ind, it means list(info(i,j,x_{ij},$\mu_{ij}^{cen_i nd}$)) Output: Pair (k3,v3): k3 is the new recalculated center c_{New} (cluster index cen_ind), v3 is list(info(i,j,x_{ij},$\mu_{ij}^{cen_i nd}$))

- Step 1: Initialize c_{New} array with the number of elements equal to the dimensions of the objects x_{ij}
- Step 2: Initialize variable totalM = 0
- Step 3: Browse list((i,j,x_{ij},$\mu_{ij}^{cen_i nd}$)
  - Step 3.1: Extract components of intensity and membership value: x_{ij},$\mu_{ij}^{cen\_ind}$
  - Step 3.2: Calculate c_{New} += x_{ij} * $\left(\mu_{ij}^{cen\_ind}\right)^{m}$
  - Step 3.3: Calculate totalM += $\left(\mu_{ij}^{cen\_ind}\right)^{m}$
- Step 4: Devive c_{New} with totalM to obtain a new center value
- Step 5: Assign k3 = c_{New}
- Step 6: Assign v3 = list((info(i,j,x_{ij},$\mu_{ij}^{cen\_ind}$))
- Step 7: Return (k3,v3)

3.2.4. *Generate the clustered image and the stage of post-clustering.* From output data of the reduce_Fuzzy function, most simply, the clustered image can be retrieved from the position information and intensity value of the cluster centers. In addition, after that, we can implement other things like data evaluation, data analysis, identification, classification, decision making, etc.

3.3. **Proving that the quality of Fuzzy C-Means and MapReduce_Fuzzy algorithms are the same.** Clause: If the same input data set and the central set are initialized, two algorithms Fuzzy C-Means and MapReduce_Fuzzy give the same clustering results. Input: Data set $X = \{x_1, x_2, ..., x_n\}$, the initialized center set $C = \{c_1, c_2, ..., c_n\}$ Proving:

- At the first loop:
  - With each data point x_i, considering membership to c_j is μ_i^j (in center set $C = \{c_1, c_2, ..., c_n\}$).
  - Comment (a): Because the distance $d(x_i, c_j)$ from the data point x_i to c_j is the same whether it is calculated in the algorithm FCM or MRF, according to

Equation (10), membership μ_i^j is the same even though this value is calculated in the algorithm FCM or MRF.
  - Comment (b): All the data points (in the input data set) and corresponding membership will be grouped according to each cluster.
  - Comment (c): From comment (b), for each cluster is represented by center cs, cs_{new} (recalculated center) calculated by the algorithm FCM or MRF is the same.
  - Comment (d): From comment (c), the output cluster center set of the first loop is the same with the algorithm FCM or MRF.
  - At the second loop: The center set is taken from the output of the first loop should be the same with the algorithm FCM or MRF.
  - With the same reasoning as in the first loop, the output center sets of the second loop are the same with the algorithms FCM or MRF.
- Reasoning as the second loop for each the next loop, we have the same results with the algorithms FCM or MRF.

Thus, the clustering results after convergence are the same with the algorithms FCM or MRF. That is thing which must be proved.

4. **Experiments.** We test the proposed algorithm MapReduce_Fuzzy (MRF) and compare to the algorithm Fuzzy C-Means (FCM). Data set used for experiments includes 3 types. The first, Landsat ETM+ images are taken in Hoa Binh area in 2001 (on 15/02/2001), including 11 pictures of districts and 1 picture of Hoa Binh province. The second, SPOT 4 images are about Hoa Binh and Son La areas with 21 pictures in 2003 and 14 pictures in 2008. The third, Quickbird images which are downloaded from model data on website: http://opticks.org. Because of the limited scope of the paper, the authors present experiments with different three input images shown in figure 5. In the experiment, we use the Spark tool to implement the algorithm MapReduce_Fuzzy using the MapReduce model.

FIGURE 5. The input images in the experiments 1, 2 and 3.

The test diagram is illustrated in Figure 6. Thus, two above algorithms have the same initialized center set.

4.1. **Experiment 1.** Input image is a Lansat image of Da Bac district, belong to Hoa binh province, with size 1596 x 1333. The clustered image is shown in fig 7.

Table 1 and Figure 8 show statistics and compare the execution time of two algorithms with Lansat image.

4.2. **Experiment 2.** Input image is a Quickbird image with size 2056 x 2065. The clustered image is shown in figure 9.

Table 2 and Figure 10 show statistics and compare the execution time of two algorithms with Quickbird image.

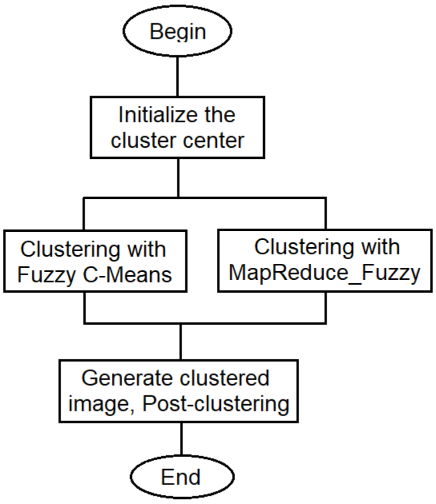

FIGURE 6. The test diagram with the algorithms Fuzzy C-Means and MapReduce_Fuzzy.

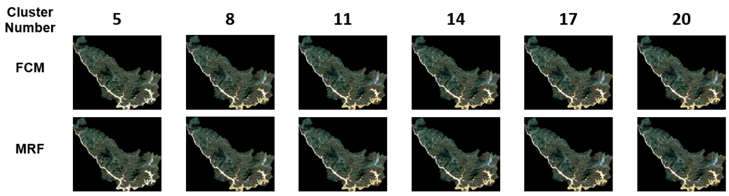

FIGURE 7. The clustered images of the Lansat image.

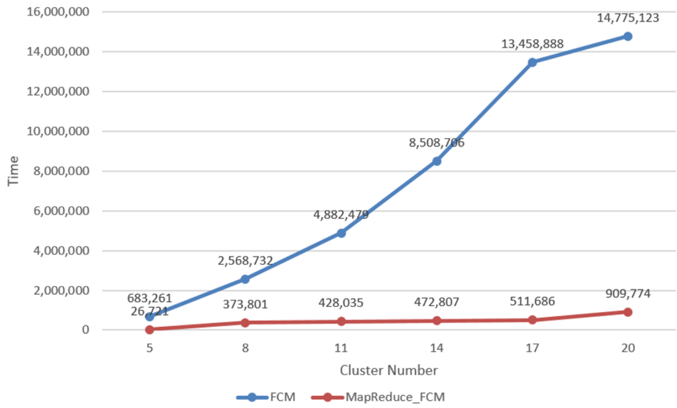

FIGURE 8. The diagram of comparing clustering time of Lansat image.

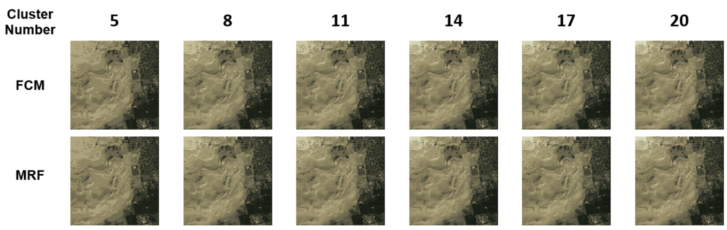

FIGURE 9. The clustered images of the Quickbird image.

Table 1. The clustering time of Lansat image.

Cluster Number	FCM	MRF
5	683,261	26,721
8	2,568,732	373,801
11	4,882,479	428,035
14	8,508,706	472,807
17	13,458,888	511,686
20	14,775,123	909,774

Table 2. The clustering time of Quickbird image.

Cluster Number	FCM	MRF
5	681,186	127,773
8	2,873,355	336,061
11	5,809,545	641,885
14	15,960,209	940,179
17	18,042,748	1,202,135
20	26,700,984	2,062,376

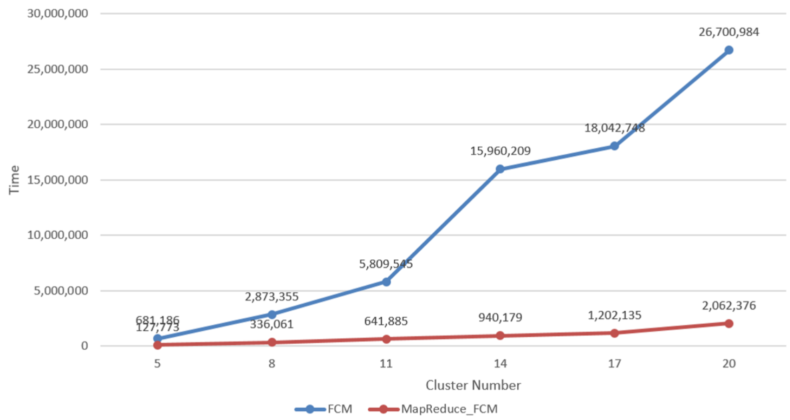

Figure 10. The diagram of comparing clustering time of Quickbird image

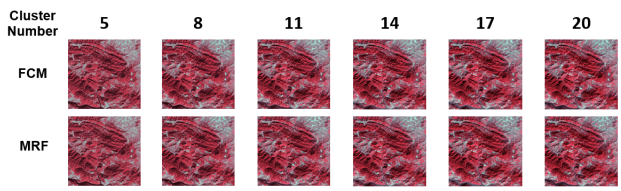

Figure 11. The clustered images of SPOT image.

4.3. **Experiment 3.** Input image is a SPOT image with size 2201 x 2101. The clustered image is shown in figure 11.

Table 3 and Figure 12 show statistics and compare the execution time of two algorithms with SPOT image.

Comments: As proved in section 3.3, the quality of the two algorithms is the same. Moreover, Table 1, 2, 3 and Figures 8, 10, 12 show that the clustering execution speed of the algorithm MapReduce_Fuzzy outperform to the algorithm Fuzzy C-Means.

TABLE 3. The clustering time of Quickbird image.

Cluster Number	FCM	MRF
5	3,441,597	179,329
8	4,786,402	819,436
11	5,797,639	854,393
14	11,689,740	1,004,062
17	13,393,833	1,110,313
20	35,209,072	2,744,031

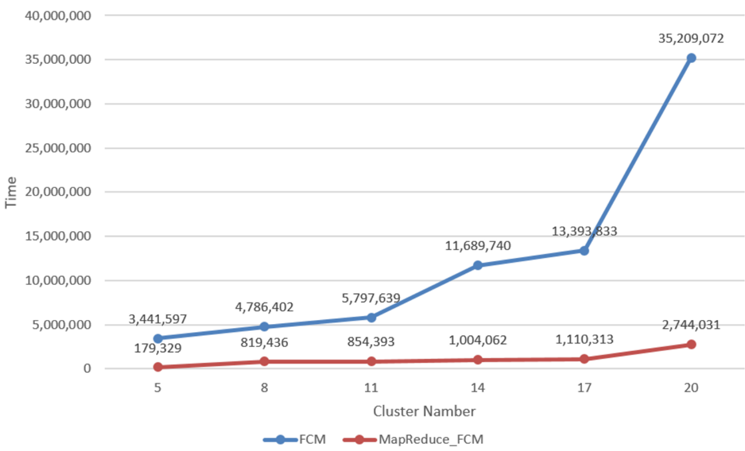

FIGURE 12. The clustered images of SPOT image.

5. **Conclusions.** In this paper, the authors proposed the new image clustering algorithm MapReduce_Fuzzy that uses the MapReduce model to improve the clustering speed of the algorithm Fuzzy C-Means. In addition, the article also presents formal representation and the detailed algorithm representations of the procedures map_Fuzzy and reduce_Fuzzy. The test results show that the algorithm MapReduce_Fuzzy gives much better clustering time compared to the algorithm Fuzzy C-Means without reducing clustering quality.

In the next study, we plan to apply the MapReduce model to other machine learning algorithms to be able to exploit, analyze and process big data efficiently.

Acknowledgment. This work is partially supported by Department of Information Technology, Thuyloi university and Institute of Information Technology, Vietnamese Academy of Science and Technology. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] Balaji T., Sumathi M., Relational features of remote sensing image classification using effective KMeans clustering, *International Journal of Advancements in Research & Technology*, vol. 2, no. 8, pp. 103-107, August-2013.

[2] Chih-Tang Chang et al, A fuzzy KMeans clustering algorithm using cluster center displacement, *Journal of Information science and Engineering 27*, pp. 995-1009, 2011.

[3] Canada Center for Remote Sensing, *Fundamentals of remote sensing*, http://www.ccrs.nrcan.gc.ca, 2008.

[4] Jeffrey Dean and Sanjay Ghemawat, MapReduce: simplified data processing on large clusters, *USENIX Association OSDI 04: 6th Symposium on Operating Systems Design and Implementation*, 2004.

[5] Hasanien A.E., A. Badr, A comparative study on digital mamography enhancement algorithms based on fuzzy theory, *Studies in Informatics and Control*, vol.12, no.1, March 2003.

[6] Herodotos Herodotou, Business intelligence and analytics: big systems for big data, *Cyprus University of Technology*, 2016.

[7] Intan aidha yusoff, Nor ashidi mat isa, Two-dimensional clustering algorithms for image segmentation, *WSEAS Transactions on Computers*, vol. 10, no.10, October 2011.

[8] J.C. BEZDEK, R. EHRLICH, W.FULL, FCM: The fuzzy c-Means clustering algorithm, *Computers & Geosciences*, vol. 10, no. 2-3, pp. 191-203, 1984.

[9] Meritxell Bach Cuadra, Jean-Philippe Thiran, Satellite image segmentation and classification, *Diploma project*, Fall 2004.

[10] Nandhini.P, A research on big data analytics security and privacy in cloud, *Data Mining, Hadoop and Mapreduce, Int. Journal of Engineering Research and Application*, 2018.

[11] Nguyen Khac Thoi et al, Remote sensing Curriculum, *Vietnam National university of Agriculture*, 2011.

[12] Shruti Dalmiya et al, Application of wavelet based KMeans algorithm in mammogram segmentatio, *International Journal of Computer Application*, vol. 52 no.15, August 2012.

[13] Tom White, *Hadoop: The Definitive Guide : The Definitive Guide*, 2009.

[14] Trung Nguyen Tu et al, Enhancing remote sensing image contrast based on combination of fuzzy logic and local spproach, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 10, no. 4, December 2019.

[15] Xiao-dong Wang, Rung-Ching Chen and Fei Yan, High-dimensional data clustering using K-means subspace feature selection, *Journal of Network Intelligence*, vol. 4, no. 3, pp. 80-87, August 2019.

[16] Xixi Dong, Yun Zhang and Shujuan Yu, An uneven clustering routing protocol based on improved K-means algorithm for wireless sensor network in coal-mine, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 10, no. 1, pp. 53-62, January 2019.