Road extraction in satellite image with fuzzy c-mean

I G A S Melati¹, N N U Januhari², A Yaputra³

¹,²,³ Study Program Information System of ITB STIKOM Bali

Email: ¹melati@stikom-bali.ac.id, ²amik@stikom-bali.ac.id, ³alberthyap@gmail.com

Abstract. Algorithm Fuzzy C-Means is a data grouping technology where the level of data presence in a class or cluster is determined by the degree of membership. The-based image segmentation clustering in this study applied the Fuzzy C Means method to obtain a more optimal cluster center. Preprocessing applied to the test image used in the proposed method was gradient operation. System evaluation was applied following two approaches, namely measuring the value validity cluster. The test results show that the FCM-Means algorithm produced better segmentation. The results of the clustering produced colors on the road. The value of each pixel was limited to the value 0 or the value 1 in the binary image. In this Fuzzy C-Means algorithm, each object/data can be a member of several clusters.

1. Introduction

In its development, road processing satellite image extraction from digital images has attracted the attention of many parties. Remote sensing is a technique used to obtain information on an area or object of interest by analyzing the data obtained using the media. With the development of technology, many segmentation algorithms have been developed, one of which is Fuzzy clustering as one of the most widely used algorithms for image segmentation. Fuzzy c-means algorithm (FCM) is the most popular fuzzy clustering algorithm and many fuzzy clustering algorithms are newly discovered from the algorithm. One of the techniques in image segmentation is clustering. Clustering is a method of classifying or partitioning data in a dataset, wherein the data broken image segmentation is the image pixels. Based image segmentation clustering can be applied to color images as well as images grayscale. The tracking of curb points in rural areas is the most difficult problem in road extraction because the edges of the points along the road are not necessarily connected within the edges of the original map. In this study, the authors carried out several sequence of operations for the extraction process of road images obtained from satellite images. The initial step taken was to change the image of a colored road or a gray road image to be binary. In this study, the authors tried to propose a system entitled Road Extraction on Satellite Imagery with Fuzzy C-Mean which is expected to contribute better in terms of standard performance such as completeness, accuracy and quality. This research proposal tried to apply the Fuzzy C Means algorithm to perform the image segmentation process based on clustering. The test images used for the testing of the developed model were some standard test images commonly used by researchers to test the developed algorithm. To determine the performance of the algorithm model developed, output the resulting was compared with the output of the standard Fuzzy C-Means algorithm. Some of the techniques used for image segmentation are edge-based and area-based. grouping is based on outlines which are divided into images based on sub-regions while segmentation based on the working area is based on the uniformity that exists in these sub-regions.
2. Literature Review
Remote sensing is the process of gathering information about an object and without a direct relationship with that object. Radar and satellites are the most common features in the remote sensing observation process. Sensing itself is very limited by techniques or methods that use electromagnetic techniques as a first feature with a detection process technique and in targeting [1]. Digital image processing is a science that discusses or studies something related to image or image quality, for example, contrast enhancement, color transformation, image restoration and image transformation from translation rotation, scale, geometric transformation. Good and optimal image selection is intended for the purpose of a good analysis in remote image processing [2], grouping an information retrieval or object description and recognizing objects in an image obtained from remote sensing, reducing the pixel value in the image is the stage of resizing or reducing data objects for the purpose of storing data, data transmission, and time in storing or processing data. The input from image processing or image is the image used while the output is the result of image processing itself [3]. Based segmentation clustering uses multidimensional data as a group of image pixels into some clustering. Broadly speaking, pixels are grouped based on similarity to pixels. Fuzzy C-means (FCM) clustering is a fuzzy grouping technique that makes data a member of objects contained in all classes or clusters that are built with different degrees or memberships between 0 and 1.

The steps for the Fuzzy C-Means algorithm are as follows [4]:

1. The **input** of data to be in cluster X, in the form of a matrix measuring n×m.
   
   by:
   n: number of data samples
   m: attributes of each data
   \( X_{ij} \): data sample \( i (i = 1,2,\ldots,n) \), the \( j \)-th \( (j = 1,2,\ldots,m) \)

2. Determine:
   Number of clusters: \( c \)
   Power: \( m \)
   Maximum iteration: \( MaxIter \)
   Minimum error expected: \( x \)
   Initial objective function: \( P_0 = 0 \)
   Initial iteration: \( t = 1 \)

3. Next calculate the number of each column (attribute)
   \[
   Q_j = \sum_{k=1}^{c} \mu_{ik}
   \]  
   (1)

4. Calculate the center of the \( k \)-th cluster: \( V_{kj} \), where \( k = 1,2,\ldots,c \); and \( j = 1,2,\ldots,m \).
   \[
   V_{kj} = \frac{\sum_{i=1}^{n} ((\mu_{ik})^m \times X_{ij})}{\sum_{i=1}^{n} (\mu_{ik})^m}
   \]  
   (2)

5. Compute the objective function in the iteration \( t \), \( P_t \).
   \[
   P_t = \sum_{i=1}^{n} \sum_{k=1}^{c} ((\mu_{ik})^m (X_{ij} - V_{kj})^2)
   \]  
   (3)
6. Calculating partition matrix changes:

\[
\mu_{ik} = \frac{\sum_{j=1}^{m} (x_{ij} - v_{kj})^{-1}}{\sum_{k=1}^{p} \sum_{j=1}^{m} (x_{ij} - v_{kj})^{-1}}
\]  

(4)

3. Research Method

The image acquisition process is the process of capturing an image of an object, in this case, a road image. In this system, the image acquisition process was used for capturing the path from a specified satellite image and then saving it as an input to the Road Extraction system on a Satellite Image that was used. In this process, the image with the object of the road was first extracted with the method Fuzzy C-Means so that later the cluster image data were obtained. The Fuzzy C-Means method is a method clustering that reallocates data into each cluster by utilizing theory Fuzzy. Image segmentation based on clustering used multidimensional data to group image pixels into several clustering. Multidimensional data in this image means the number of attributes or components that make up an image, for example, a grayscale image has one dimension, an RGB image has three dimensions, and so on. In general, pixels are clustered based on the proximity of the pixels. The success of-based segmentation process clustering is determined by the success in grouping adjacent features into one cluster. One excellent method used for image segmentation is Fuzzy C-Means clustering. Fuzzy C-Means is a k-means algorithm that is improvised with Fuzzy Set Theory by applying a degree of membership, where one image pixel can be owned by several clusters. This "soft" clustering provides a more precise computation in determining the membership of the cluster. The number of clusters here is the number of clusters or colour groups desired in the segmented image. The FCM algorithms were tested using different cluster number parameters to determine the optimal number of clusters from each test image. The number of clusters here is the number of clusters or colour groups desired in the segmented image. The standard FCM algorithm was tested using different cluster number parameters to determine the optimal number of clusters from each test image.

3.1 C-Means Clustering Extraction

The basic concept of FCM is to define a cluster which marked each location in a cluster object. In a cluster of each cluster membership has a cluster point simultaneously. The existing cluster have the same location as the previous cluster [5]. This algorithm step was based on the minimum of functions and objectives that described the distance and certain data points to the cluster which are weighted by the degree of membership of the data points. The FCM function in road extraction can be described in the following diagram. Each cluster has a degree of membership owned by each cluster. Cluster repetition of each data set in the membership of each subject went to the appropriate location. This repetition was based on a minimum function of a distance from a particular data point to a cluster assessed by the degree of membership of the data object from the fuzzy set. This iteration was based on minimizing the objective function describing the distance from a given data point to a cluster weighted by the degree of membership of that data point. The function of FCM on road images can be described in the following flowchart. At initial conditions, the of cluster this is still inaccurate. Each data has a degree of membership for each cluster. By the cluster repeating and membership value of each data set, it can be seen that the cluster went to the right location. This loop was based on minimizing the objective function describing the distance from a given data point to the cluster weighted by the degree of data point membership of the set fuzzy. In Fuzzy C-Means each object can be a member of multiple clusters. Fuzzy C-Means Algorithm places image pixels into each cluster using membership fuzzy. \( X = (x_1, x_2, \ldots, x_N) \) shows an image with \( N \) pixels that will be partitioned into \( c \) clusters. This algorithm is an iterative optimization that minimizes the cost function, where the cost function used in this study is the objective function of the Index.
4. Main Result
This system interface was used to get road extraction results. This application was built using the C# programming language in the Visual Studio 2008 editor. The application "road extraction using the FCM method" was the main interface of this study.
The test image was displayed on the main panel which can then be processed using the standard FCM method or the FCM-CSO. To load or insert a test image in this application, select "Open" which is on the "File" menu, then proceed by selecting the desired image file where the types of files that can be loaded here are image files of type JPG, PNG, TIF, TIFF, BMP, and GIF.

Image acquisition process is a process of making (Capturing) image of an object, in this case, the image of the road, which in this system is the capture of Peng's way satellite images was determined and then stored as the input of the system extraction Jalan On with Satellite Imagery Fuzzy C-Mean
Figure 4. Clustering 1.

Figure 5. Number of clusters.
Figure 6. Number of clusters.

The results from taking the road image were then stored in a database, in this case, the database was stored on a computer hard drive. The reason for the road image database diagram is to be used in the future for image research needs that have already been taken. The data from the database for the road image was then needed and directed to the extraction process using Fuzzy C-Means.

Figure 7. Coloring Layer.
In Figures 7 and 8, the process of coloring the border area, values of all transformed pixels were identified using the modulus operation performed on the boundary values pixel. Furthermore, the application operation scared of morphology closed on the image was out to fill small holes in objects and combine adjacent objects and in general smooth the boundaries of the objects without changing the object area significantly.

Table 1. Clustering.

| The Number of Cluster | The Number of Iteration | Estimated Time |
|-----------------------|-------------------------|----------------|
| 2                     | 10                      | 00:00:07       |
| 3                     | 20                      | 00:00:21       |
| 4                     | 100                     | 00:02:40       |
| 5                     | 17                      | 00:08:78       |
| Average               | 67                      | 746 second     |

Based on the results of this trial, it can be seen that the most optimal number of clusters is 2 clusters, which produces the smallest XB Index value.

5. Conclusion
The results of satellite image processing on the road showed a change in the curb lines with colour. Clustering was used to group road data and improve the image to accentuate the character of the image. Filter selection was a step that reduces unnecessary small objects to be processed into the system. The result of blurring the image with the Gaussian function was widely used to reduce objects (blur). Make a gaussian filter according to the given N value and perform image with clustering at the process stage $D_{1x}$, $D_{1y}$ results from the previous process and the same process is repeated for all objects in the
result. The image generated a binary image in the process of coloring the border area values of the changing pixels identified using the modulus operation performed at the boundary values pixel. Furthermore, the application operation scared of morphology closed on the image was out to fill small holes in objects and combine adjacent objects and in general smooth the boundaries of the objects without changing the object area significantly.

References
[1] Ambarwati A, Passarella R and Sutarno 2016 Segmentasi Citra Digital Menggunakan Thresholding Otsu untuk Analisa Perbandingan Deteksi Tepi Annu. Res. Semin. 2 216–226
[2] Wijaya N 2015 Deteksi Perubahan Penggunaan Lahan Dengan Citra Landsat Dan Sistem Informasi Geografis: Studi Kasus Di Wilayah Metropolitan Bandung, Indonesia Geopanning J. Geomatics Plan. 2 82–92 doi: 10.14710/geopanning.2.2.82-92
[3] Arsy L, Nurhayati O D and Martono K T 2016 Aplikasi Pengolahan Citra Digital Meat Detection Dengan Metode Segmentasi K-Mean Clustering Berbasis OpenCV Dan Eclipse J. Teknol. dan Sist. Komput. 4 322 doi: 10.14710/jtsiskom.4.2.2016.322-332
[4] Sianturi C J M 2015 Analisis Segmentasi Citra USG Hati Menggunakan Metode Fuzzy C-Mean Citec J. 2 256–264
[5] Hardiyanto I, Purwananto Y and Soelaiman R 2012 Implementasi Segmentasi Citra dengan Menggunakan Metode Generalized Fuzzy C- Means Clustering Algorithm with Improved Fuzzy Partitions Tek. Pomits 1 1–5