



Spatial Modification in the Parameters of Mountain Image Clustering Algorithm

Nahla Ibraheem Jabbar

Authors affiliations:

1) Department of Electrochemical Eng., College of Engineering, Babylon University, Babylon-Iraq.
az2008na@gmail.com

Abstract

Our proposed method used to overcome the drawbacks of computing values parameters in the mountain algorithm to image clustering. All existing clustering algorithms are required values of parameters to starting the clustering process such as these algorithms have a big problem in computing parameters. One of the famous clustering is a mountain algorithm that gives expected number of clusters, we presented in this paper a new modification of mountain clustering called Spatial Modification in the Parameters of Mountain Image Clustering Algorithm. This modification in the spatial information of image by taking a window mask for each center pixel value to compute distance between pixel and neighborhood for estimation the values of parameters σ , β that gives a potential optimum number of clusters requiring in image segmentation process. Our experiments show ability the proposed algorithm in image brain segmentation with a quality in the large data sets

Keywords: Clustering, Mountain Algorithm, Image Segmentation and Spatial Information.

التحوير المكاني في معلمات خوارزمية الجبل لتقسيم الصورة

نهلة ابراهيم جبار

الخلاصة:

استخدام الطريقة المقترحة للتغلب على عيوب حساب قيم المعلمات في خوارزمية الجبل لتقسيم الصورة. يتطلب من جميع الخوارزميات التصنيف لبدء عملية التصنيف قيم المعلمات. والخوارزميات لديها مشاكل كبيرة في حساب هذه قيم المعلمات. من احدى الخوارزميات الشهيرة هي خوارزمية الجبل التي تعطي العدد المتوقع للتصنيف. قدمنا في هذا البحث على تعديل جديد من تجميع جبل يسمى تعديل المكاني في معلمات من جبل تقسيم الصورة. هذا التعديل يعتمد على المعلومات المكانية للصورة عن طريق أخذ قناع نافذة لكل قيمة بكسل، وحساب المسافة بين البيكسل والمتجاورات. والذي يعطي العدد الأمثل بمعلمات المطلوبة σ و β في عملية تجزئة الصورة. تظهر تجاربنا قدرة الخوارزمية المقترحة في صورته تجزئة الدماغ مع جودة في التصنيف في التصنيف مجموعات كبيرة من البيانات.

1. Introduction

Clustering is the unsupervised classification of patterns. Clustering is a process of partitioning or grouping a given set of unlabeled patterns into a number of clusters such that similar patterns are assigned to one cluster. Each pattern can be represented by a vector having many parameters or attributes. Fundamental to the use of any clustering technique is the computation of a measure of similarity or distance between the respective patterns [1]. Clustering is useful in pattern-analysis, grouping, decision-making, machine-learning, data mining, document retrieval, image segmentation, and pattern classification. Today many data clustering algorithms are being used for segmenting images. They are termed as unsupervised methods for segmentation of images. In such techniques, image is separated into a set of disjoint regions with each region associated with one of the finite number of classes that are characterized

by distinct parameters. Therefore, still date many types of segmentation techniques have been developed and many data clustering techniques are being used for segmentation of images [2]. One of the clustering techniques is a mountain clustering. It is proposed by Yager and Filev [3]. Hung and others proposed a feature weighted in mountain clustering that application in color image segmentation [4]. In the paper [5] proposed Improved Mountain Clustering version-2 (IMC-2) based medical image segmentation. The proposed technique is a more powerful approach for medical image based diagnosing diseases like brain tumor, tooth decay, lung cancer. In general mountain method is a grid based three-step procedure. In the first step the hyperspace is discretized with a certain resolution in each dimension so that grid points are obtained. The second step uses the data set to construct the mountain function around all grid points. The third step generates the cluster centers by an iterative destruction of mountain function.



2. Mountain Clustering Method

The mountain clustering method is a relatively simple and effective approach to approximate estimation of cluster centers on the basis of a density measure called the mountain function. The mountain clustering approach is a simple way to find cluster centers based on a density measure called the mountain function. This method is a simple way to find approximate cluster centers. The first step in mountain clustering involves forming a grid on the data space, where the intersections of the grid lines constitute the potential cluster centers, denoted as a set V . The second step entails constructing a mountain function representing a data density measure. The height of the mountain function at point $v \in V$ is equal to [6].

$$m(v) = \sum_{i=1}^n \exp\left(-\frac{\|v - x_i\|^2}{2\sigma^2}\right) \dots \dots \dots (1)$$

where x_i data point and σ is an application specific constant. This equation states that the data density measure at a point v is affected by all the points x_i in the data set, and this density measure is inversely proportional to x_i and the point under consideration. The constant σ the distance between the data points determines the height as well as the smoothness. The third step involves selecting the cluster centers by sequentially destructing the mountain c is determined by electing the point with the greatest density function measure. Obtaining the next cluster center requires eliminating the effect of the first cluster. This is done by revising mountain function: a new mountain function is formed by subtracting a scaled Gaussian function centered at c_1 [1].

$$m_{new}(v) = m(v) - m(c_1) \exp\left(-\frac{\|v - c_1\|^2}{2\beta^2}\right) \dots (2)$$

The subtracted amount eliminates the effect of the first cluster. Note that after subtraction, the new mountain function $m(v)$ reduces to zero at $v = c_1$. After subtraction, the second cluster center is selected as the point having the greatest value for the new mountain function. This process continues until a sufficient number of cluster centers.

3. Modified Mountain Clustering Algorithm

Conventional mountain algorithm classifies pixel data based their feature space distribution without considering the spatial interactions between neighboring pixels. The clustering result will be influenced by the coordinate system chosen when the coordinates are used as features. The value of the mountain function can be related to the density data points in neighborhood of the node. The higher mountain function value the larger potential ability. This potential depends on the grid node to be cluster center [7].

Spatial relationship in an image are often of information between regions. The set of feature vectors associated with an image of different sizes by, where is the p-dimensional. The feature vector describes the properties or attributes to the pixel at location (i, j) [8]. For a pixel at location gives the

Euclidean distance between feature vectors and its neighbors by using a window or mask is used to compute the difference between pixel value of the center with neighbors' pixels. The size of window is 3x3, 5x5 and 7x7, the higher dimension of the mask gives higher number of pixels in the mask and faster in the computation [9]. In our experiments the size of window is 3x3. The proposed algorithm takes consider the distance between center pixel and the neighbors pixels which that used to compute the parameter θ also this algorithm deals the problem parameter β by weighting the difference distance center pixels with neighbored and the average distance in all image ,after applied these modification in in mountain algorithm ,the new algorithm has been as flows:-

4. Algorithm of Spatial Modification Parameters in Mountain Image Clustering

Step 1: Normalize all data points(pixels)

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \dots \dots \dots (3)$$

$$x_i = \{x_1, x_2, x_3, \dots \dots x_n\}$$

n is the total number of data points

Step 2: Determine the threshold value d_1 for each window. d_1 is positive defining the difference neighborhood of data point [1].

$$d_1 = \|x_i - x_j\| \dots \dots \dots (4)$$

Where d_1 for each data point x_i

$$d(x_i, x_j) = d(x_i - 1, x_j - 1) + d(x_i - 1, x_j) + d(x_i - 1, x_j + 1) + d(x_i, x_j - 1) + d(x_i, x_j) + d(x_i, x_j + 1) + d(x_i + 1, x_j - 1) + d(x_i + 1, x_j) + d(x_i + 1, x_j + 1)$$

The equation above for the calculation of values between the center pixel and the neighborhood pixels of the three (3 by 3) masks using spatial convolution process. Every pixel in the input image is evaluated with its eight neighbors, using each of the three masks shown in Fig. (1).

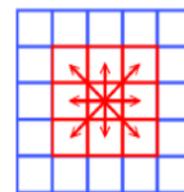


Figure (1): window size (3 by 3)

$$d_2 = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \frac{1}{8} \sum_{l_1=1}^3 \sum_{l_2=1}^3 d\{(i, j), (i + l_1, j + l_2)\} \dots \dots \dots (5)$$



Step 3: Calculate the potential value of each point p using mountain function, which is a function of distance

$$p_i = \sum_{j=1}^n \exp - \left(\frac{d_1}{d_2} \right) \dots \dots \dots (6)$$

Step 4: Calculate the highest value of mountain function

$$c_1 = \max(p_i) \dots \dots \dots (7)$$

Step 5: Assign those data points to the first cluster whose Euclidean distance from the first cluster center is less than a threshold

$$\text{distance}(x_i, c) \leq d_1 \dots \dots (8)$$

Step 6: Remove all those data points from the total data set which are assigned to the cluster formed.

Step 7: Repeat step2 to 6 for the remaining data to make clusters.

5. Result and Discussion

The algorithm described above has been applied to the different types of images (gray and color images) .and best results has been taken by trial and error. In first type of image is used a gray medical image of brain. The results are shown in figures 2(a) and (b). show results of Spatial Modification Parameter brain image in Mountain Image Clustering with $c=3$ where c is the number of clusters. The successful of proposed algorithm in Fig 3(a) the input is a color MRI brain tumor image (RGB) after applied algorithm the results as shown in Fig.3(b).

The number of clusters is given by the user as 5 color brain images. Fig.3(c) shows results of FCM clustering [11]. During the application the distance value of $d1$ was taken as 0.15. Though the effect of $d1$ is on mountain function makes the two potential values closer or farther numerically. The effect of $d2$ is very small or very large will in the number of significant clusters. A large $d2$ makes the density function will be affected account all the data points in the data space. So, a value between 0.4 and 0.7 should be suitable for radius of neighborhood. If $d2$ is very small the density function will not take into account the effect of neighboring data points in the data space. As seen in the Table (1).

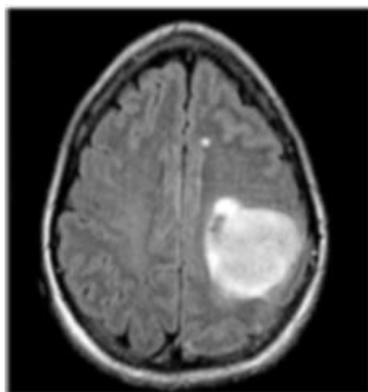


Figure (2a): An original image

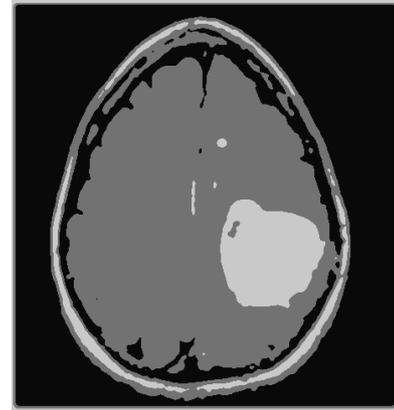


Figure (2b): Results of segmentation $c=3$ where c is the number of clusters.

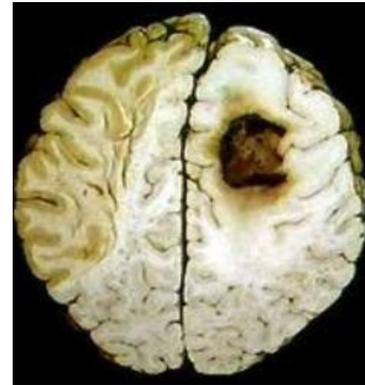


Figure (3a): An original MRI brain tumor image (RGB).

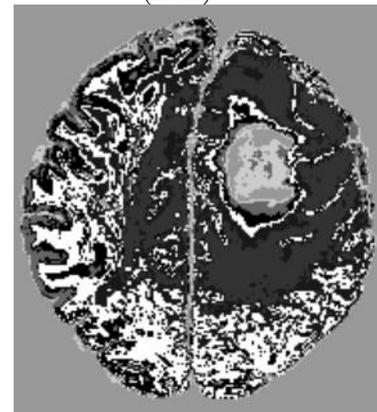


Figure (3b): Results of segmentation

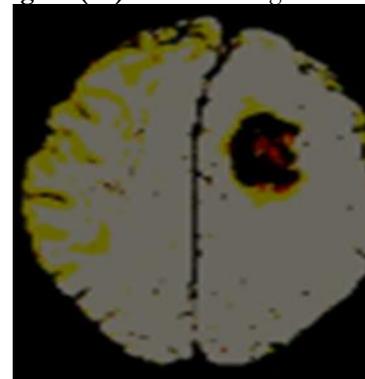


Figure (3c): shows results of FCM segmentation $c=5$

According to the implementation result of Spatial Modification Parameters in Mountain



Image Clustering and the results implementation of Fuzzy C-mean (FCM) algorithm [11]. The time execution the

proposed algorithm more than FCM. We can see that in Table.2.

Table (1): Spatial Modification Parameters in Mountain Image Clustering

Performance Measure	Neighborhood distance								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Accuracy	55.0%	58.0%	58.0%	75.0%	75.0%	75.0%	75.0%	75.0%	58.0%

Table 2. Performance in time execution and recall between two algorithms

Type of Clustering	Recall	Execution Time(sec)
Spatial Modification Parameters in Mountain Image Clustering	94	70
Fuzzy C-mean Clustering	60	20

6. Conclusions

Mountain clustering is suitable only for problems with two or three clusters. The modification of algorithm to update automatically during computing the value of parameter β in origin algorithm is modified according the distance between input and neighborhood his paper presents a new clustering approach called Spatial Modification Parameters in Mountain Image Clustering. The main concept is that we define a mountain function at each element of the data set, which forms a potential cluster, and calculate the strength of this function as a function of distance of neighboring elements. On the basis of the strength it is declared as a cluster and the

Mountain clustering is suitable only for problems with two or three clusters. The modification of algorithm to update automatically during computing the value of parameter β in origin algorithm is modified according the distance between input and neighborhood his paper presents a new clustering approach called Spatial Modification Parameters in Mountain Image Clustering. The main concept is that we define a mountain function at each element of the data set, which forms a potential cluster, and calculate the strength of this function as a function of distance of neighboring elements. On the basis of the strength it is declared as a cluster and the effect of this is removed on all other data elements. Next, another element is chosen as next potential cluster center. This procedure is repeated until a validity criterion comprising a ratio of compactness of the clusters to the separation among the clusters is violated. The results are comparable to the results of fuzzy C-means technique but Spatial Modification Parameters in Mountain Image Clustering computationally much more efficient. In this brain tumors image have been clustered better in the proposed algorithm. The execution time of Spatial Modification Parameters in Mountain more than fuzzy C-means technique. Regarding the number of tumor pixels gave a better result in Spatial Modification Parameters in Mountain Image Clustering than the other methods. The system can be extended to detect tumors on in other views of the brain also, this system considers the color in brain image. The incorporation of spatial information can resolve this ambiguity and yields better classification result. We consider a 3x3 image window.

Finally, when evaluated the algorithm with other different algorithms number of error measures are discussed and calculated. For the parameters measured, it was concluded that single algorithm will never work better for all fields are based on these numerical values, objectivity of various algorithms is defined and the performance it gives for particular application. In the future work the performance analysis will be based on some newer optimization techniques as well as the algorithms and comparison will be extended to wide range of applications.

7. References

- [1] Veronica S. Moertini, "Introduction to five data clustering algorithm", Integral, 7(2), Octobr 2002.
- [2] Feng Ge, Song Wang and, Tiecheng Liu, "New Benchmark for image segmentation evaluation", Journal of Soft Electronic Imaging, 16(3), 033011, pp.033011-1-16, Jul-sep 2007.
- [3] Yager and Filev Duda RO, Hart PE (1973) Pattern classification and scene analysis. Wiley, RO, Hart PE 1973, New York.
- [4] Wen-Liang Hung¹, Miin-Shen Yang², Jian Yu³, Chao-Ming Hwang⁴, "Feature-Weighted Mountain Method with Its Application to Color Image Segmentation", International Journal of Computational Intelligence Systems, 4(5) Sep. 2011.
- [5] Verma, N.K, "Medical Image Segmentation Using Improved Mountain Clustering Technique Version-2", Information Technology: New Generations (ITNG), 2010 Seventh International Conference on 12-14 April 2010
- [6] Veronica S.Moertini,"Introduction to Five Data Clustering Algorithms", INTEGRAL 7(2)Oct. 2002.
- [7] Ronald R. Yager and Dimitar P Filev," Approximate Clustering Via the Mountain Method", Machine intelligence Institute. Iona College.
- [8] Lei jiang, Wenhui Yang, 2003 "A modified Fuzzy C-means algorithm for Magnetic Resonance
- [9] Gonzalez, R.C. and R.E. Woods, "Digital image processing, Pearson Education", 2002.
- [10] Fu, K.S. and J.K. Mui, 1981." A survey on image segmentation", Pattern Recognition.
- [11] Juraj Horvath, 2006 "Image Segmentation using Fuzzy C-means" SAMI 2006.