A ROBUST METHOD FOR REDUCING IMAGE NOISE IN MICROARRAY IMAGES

1Durga Prasad Kondisetty, 2Dr. Mohammed Ali Hussain,
1Research Scholar, Dept. of Computer Science, Bharathiar University, Tamilnadu, India.
2Professor, Dept. of Computer Science & Engineering, KL University, Vijayawada, AP, India.
dpkondisetty@gmail.com, alihussain.phd@gmail.com

Abstract: An image robust segmentation algorithm called Extended Self organizing maps (ESOM) is explained here which works (preprocessing) on the image is to reduce noise effect and then apply SOM algorithm for segmentation of an image. The preprocessing of an image is connections by all the eight surrounding pixels in the direct study of an image pixel under consideration. The goals of the proposed algorithm are: (i) by compare with other techniques, it takes minimum execution time to do. (ii) By compare with other techniques, it provides more homogeneous regions. (iii) This will remove spots of noisy and is minimum sensitive to noise. This technique is a robust method for reducing noise from image segmentation with minimum computation time and minimum convergence rate by compared to any other techniques.

Keywords: Self organizing maps, Robust Segmentation and Clustering.

I. Introduction

A. In the image segmentation the image analysis and computer vision sector the image segmentation plays an important role. Image segmentation [1] is the process of partitioning or segregating the digital image into multiple segments. The main goal of segmentation is to change the representation of an image or to simplify into set of disjoint regions with uniform and homogeneous attributes such as intensity, color, tone etc. that is more essential and simple to analyze. Out of four major categories such as clustering the image, thresholding, edge detection, and segmentation in image processing [18], the two are very frequently similar to each other: those are clustering the image and segmenting the image. The result of image segmentation is a group or set of segments that collectively cover the total image representation, or a set of contours extracted from the image.

The objects are blurred and distorted due to the imaging acquisition process in between the boundaries of color images. In additional, object definitions are not always crisp and knowledge about the objects in a scene may be vague. Self Organizing Maps, or SOMs provide a way of multidimensional data representation in much lower dimensional spaces - usually one or two dimensions. They were invented by Teuvo Kohonen, a professor of the Academy of Finland also known as Kohonen Self Organizing Feature Maps.

Vector quantization is the process of reducing the dimensionality of vectors, is essentially a data compression technique. In other words, the Kohonen technique creates a network that stores information in such a way that any topological relationships within the training set are maintained. Self Organizing Maps clustering models have give the solution for the color clustering problem. Such classified models have to be analysis the clustering and classifier the designs in the fields like astrophysics, topography, therapeutic imaging used with much number of features and clusters.

This paper is as follows: Section II, review Self Organizing Maps (SOM). Section III explains the proposed Extended Self Organizing Maps (ESOM) algorithm. Section IV experimental results and Section V contains concluding remarks.

II. Background Information

Here we briefly discusses the Self Organizing Maps (SOM) [10] and Intuitionist Self Organizing Maps (ISOM) algorithms. In this paper, the data-set is denoted by ‘W’, where $W = \{w_1, w_2, w_3, \ldots, w_n\}$ specifying an image with ‘n’ pixels in M-dimensional space to be partitioned into ‘c’ clusters and centroids of clusters are denoted by $v_i$ and $w_k$ is the distance between $w_k$ and $v_i$.

A. The Self Organizing Maps Algorithm

The goal of self-organizing map is to cause different parts of the cerebral cortex in the human brain. The weights of the neurons are initialized either to sampled evenly from the subspace spanned by the two largest principal component eigenvectors or small random values.
Algorithm
Step 1: Randomize the map's nodes' weight vectors.
Step 2: Get an input vector P(t).
Step 3: Traverse each node in the network map
1. By using the Euclidean distance formula to find the equality between the input vector and the map's node's weight vector
2. Track the node that produces the lowest distance (this is the best matching unit (BMU))
Step 4: Update all the nodes in surrounding of the BMU (including the BMU) by pulling them closer to the input vector
\[ W_v(s + 1) = W_v(s) + \alpha(u, v, s) \cdot \alpha(s) \cdot (P(t) - W_v(s)) \]
Step 5: Increase s and repeat from step 2 while s < λ

Where s is the current iteration, λ is the iteration limit, t is the index of the target input data vector in the input data set P, P(t) is a target input data vector, v is the index of the node in the map, W_v is the current weight vector of node v, u is the index of the best matching unit (BMU) in the map, o is a restraint due to distance from BMU usually called the neighborhood function, and \( \alpha(s) \) is a learning restraint due to iteration progress.

B. SOM Clustering

In agglomerative clustering, the SOM neighborhood relation can be used to constrain the possible merges in the dendrogram construction. In other words, “knowledge of interpolating units can be utilized both in partitive and agglomerative clustering by excluding them from the analysis” [2]. The interpolative units form borders on the map that the construction of the dendrogram when partitive and agglomerative clustering used together with the neighborhood constraint in agglomerative clustering”.

When number of samples in a cluster is small the distance measures used in the gap criterion are increase very sensitive to noise [20]. However, the problem can be alleviated somewhat by requiring a wider gap if the sample size is small. Let x be a random sample of size N, whose mean can be estimated as:

\[ \mu = \frac{E\{x\}}{N} = \sum x/N \]

It is easy to show that its variance is

\[ \text{Var}\{\mu\} = \frac{\text{Var}\{x\}}{N}. \]

In many distribution models like the gamma distribution, the variance is directly proportional to expectation

\[ \text{Var}\{x\} = \alpha \cdot E\{x\}. \]

Therefore, let

\[ \text{Var}\{\mu\} = \alpha \cdot E\{x\}/N \]

Now, let x be the distance between each sample and its nearest neighbor in cluster Q_k of size N_k samples. A robust upper limit for nearest neighbor distance S_{nn}(Q_k) can be acquired by

\[ S_{\text{nn}}(Q_k) = E\{x\} + \alpha_2 \cdot \text{Var}\{x\} = E\{x\} + \alpha_2 \cdot \alpha_1 \cdot E\{x\}/N_k = E\{x\} \cdot (1 + \alpha/N_k) \]

In order to make also the between-clusters distance d_{Q_k, Q_l} more robust, the mean of four shortest distances between the samples can be used instead of only the shortest distance.

III. The Proposed Technique, Extended Self Organizing Maps (ESOM)

SOM is a valuable segmentation technique when there is no noise in the image. In this view, we are going to proposing a very clean and effective method, which will preprocess the noisy pixels in the image and then will apply the conventional SOM algorithm for segmenting the image.

The pixels in the immediate neighborhood possess the same feature data i.e., in other words the pixels in an image are highly correlated. “In this paper we are proposing a technique, by processing every pixel of an image based upon its immediate neighborhood pixels is to recreate the value of noisy image pixels.” The surrounding structure is considered here as shown in Fig. 1. The value of each and every pixel in the image is influenced by its direct next eight surrounding pixels. In the consideration of an image in RGB, let size of an image is H*V*N, where H is the number of rows, V is the number of columns, and N is 3 (for color image). So every pixel will be calculate by using following equation:

\[ \text{image}(i, j, n) = \sum_{n=1}^{N} \frac{1}{\text{image}(i, j, n)} \]

International Journal of Pure and Applied Mathematics Special Issue

442
Where \( K = 8 \), denoting 8 surrounding points of a study pixel. \( \text{pixel (p, q, r)} = \text{pixel intensity value at p, q, r location} \). \( \text{pixel (i, j, n)} = \text{study pixel intensity value} \).

**Figure 1. Study pixel structure**

In the homogeneous region, this will give the clustering result remains unchanged and fortify the same characteristics. However, this formula will try to regenerate the value of original pixel for a pixel noisy. As a result, from noisy regions the misclassified pixels can easily be corrected. Then after processing the pixels noisy, SOM is used to do segmenting the image. Henceforth, we are processing the image in only once before segmentation, which won’t increase the complexity of compare to Spatial SOM and other clustering algorithms for segmentation. The other methods mostly changed the standard SOM equations and introduced new methods which will calculate for every pass of the SOM algorithm that means they are extremely computationally in-depth so the strength of the extended technique is its less execution time, which was approximately similar to the conventional SOM method. Furthermore, it does not modify SOM algorithm, so does not lost the continuity of SOM function which was given in many applications like MATLAB. Different types of steps involved in the proposed technique used for image segmentation are shown in Fig. 2.

![Figure 2. Steps used for segmentation using proposed algorithm along with input and output values.](image)

**IV. Results**

In this following section, the experimental results are presented to compare between the segmentation of performance of ESOM with SOM and CSOM. Two types of images are used: (1) Brain Cancer image (2) Kidney image. For all data-sets, we are assuming the following computational protocols: \( \epsilon = 0.00001 \), Total number of iterations = 100. We chose \( m=2 \), which is a default value for SOM clustering. Experiments were implemented and simulated using MATLAB Version 7.8.

**A. Brain Cancer Image**

A grayscale of brain cancer image representing two intensity levels were used. It was corrupted with Gaussian noise. The intensity levels lies between two classes 0 and 255. Fig. 3(a) & 3(b) shows original and noisy image. Fig. 3(c) & 3(d) shows the segmentation result of SOM, ESOM respectively. Although SOM is also strongly segmented the image into the desired set of regions but did not totally recover it from the noise whereas ESOM will totally recover the image from noise. In addition, we required several clustering validity functions that are to analyses of the qualitative and quantitative evaluation. Table I shows performance comparison of SOM, CSOM, and our proposed ESOM in terms of cluster validity functions.

**B. Kidney Image**

A real kidney image consisting of two classes that are considered and corrupted with Gaussian noise to show the performance of technique. The kidneys are separated from its background values using all the segmenting methods. Fig. 4(a)-(b) shows original image and its noise respectively. Fig. 4(c)-(f) shows values of segmentation result. It is observed that ESOM is able to retain the boundary of noise more effectively by compared with other methods. TABLE-I shows that Performance of algorithms based on execution time.

![Figure 2. Steps used for segmentation using proposed algorithm along with input and output values.](image)
V. Conclusion

Images especially digital images contain noise and considerable skepticism. In General, SOM is a popular segmentation technique for digital images. Whereas it is intensity based clustering algorithm which is not robust on noisy images. The pixels of images are correlated and this spatial data is an important characteristic for digital image segmentation which is not yet utilized in SOM. Although, some extended algorithms based upon SOM have been implemented to overcome this issue by incorporating spatial information either in the objective or in the membership function of SOM algorithm. But the other functions have increased the complexity of the original algorithm by introducing new functions that are calculated for every going pass of SOM algorithm.

Here, we proposed a technique; an extended SOM (ESOM) by processing every pixel of an image based upon its immediate neighborhood pixels is to recreate the value of noisy image pixels. The Preprocessing of an image was influenced by the all its direct surrounding pixels of every pixel in an image was consideration. Its pros were its performance in least execution time and fast convergence rate as compared to other segmentation techniques. It was tested on noisy synthetic images like brain tumor and kidney images and evaluated quantitatively with different cluster validity functions. Results displayed that the performance of proposed technique is the best over SOM.

References


