

Segmentation of Medical Image using Clustering and Watershed Algorithms

¹M.C. Jobin Christ and ²R.M.S. Parvathi

¹Department of Biomedical Engineering,
Adhiyamaan College of Engineering, MGR Nagar, Hosur, India
²Sengunthar College of Engineering, Tiruchengode, India

Abstract: Problem statement: Segmentation plays an important role in medical imaging. Segmentation of an image is the division or separation of the image into dissimilar regions of similar attribute. In this study we proposed a methodology that integrates clustering algorithm and marker controlled watershed segmentation algorithm for medical image segmentation. The use of the conservative watershed algorithm for medical image analysis is pervasive because of its advantages, such as always being able to construct an entire division of the image. On the other hand, its disadvantages include over segmentation and sensitivity to false edges. **Approach:** In this study we proposed a methodology that integrates K-Means clustering with marker controlled watershed segmentation algorithm and integrates Fuzzy C-Means clustering with marker controlled watershed segmentation algorithm separately for medical image segmentation. The Clustering algorithms are unsupervised learning algorithms, while the marker controlled watershed segmentation algorithm makes use of automated thresholding on the gradient magnitude map and post-segmentation merging on the initial partitions to reduce the number of false edges and over-segmentation. **Results:** In this study, we compared K-means clustering and marker controlled watershed algorithm with Fuzzy C-means clustering and marker controlled watershed algorithm. And also we showed that our proposed method produced segmentation maps which gave fewer partitions than the segmentation maps produced by the conservative watershed algorithm. **Conclusion:** Integration of K-means clustering with marker controlled watershed algorithm gave better segmentation than integration of Fuzzy C-means clustering with marker controlled watershed algorithm. By reducing the amount of over segmentation, we obtained a segmentation map which is more diplomats of the several anatomies in the medical images.

Key words: Clustering, segmentation, medical image

INTRODUCTION

Image segmentation is a vital method for most medical image analysis tasks. Segmentation is an important process to extract information from complex medical images. Segmentation has wide application in medical field (Christ and Parvathi, 2011; Gonzalez *et al.*, 2004; Jayaraman, 2009; Murugavalli and Rajamani, 2007; Ng *et al.*, 2006; Shanmugam *et al.*, 2011; Velmurugan and Santhanam, 2010). Having good segmentations will help clinicians and patients as they provide vital information for 3-D visualization, surgical planning and early disease recognition. Image segmentation algorithms are classified into two types, supervised and unsupervised. Unsupervised algorithms are fully automatic and partition the regions in feature space with high density. The different unsupervised algorithms are Feature-Space Based Techniques,

Clustering (K-means algorithm, C-means algorithm, E-means algorithm), Histogram thresholding, Image-Domain or Region Based Techniques (Split-and-merge techniques, Region growing techniques, Neural-network based techniques, Edge Detection Technique), Fuzzy Techniques. The watershed segmentation technique has been widely used in medical image segmentation. Watershed transform is used to segment gray matter, white matter and cerebrospinal fluid from Magnetic Resonance (MR) brain images. The method originated from mathematical morphology that deals with the topographic representation of an image. Watersheds are one of the typical regions in the field of topography. A drop of the water falling it flows down until it reaches the bottom of the region. Monochrome image is considered to be a height surface in which high-altitude pixels correspond to ridges and low-altitude pixels correspond to valleys. This suggestion

Corresponding Author: M.C. Jobin Christ, Department of Biomedical Engineering, Adhiyamaan College of Engineering, Dr.MGR Nagar, Hosur, India

says if we have a minima point, by falling water, region and the frontier can be achieved. Watershed uses image gradient to initial point and region can get by region growing. The accretion of water in the neighborhood of local minima is called a catchment basin. Watershed refers to a ridge that divides areas shattered by different river systems. A catchment basin is the environmental area draining into a river or reservoir. If we consider that bright areas are high and dark areas are low, then it might look like the plane. With planes, it is natural to think in terms of catchment basins and watershed lines. Two approaches are there to find watershed of an image:

- Rainfall approach
- Flooding approach

In rainfall approach, local minima are found all through the image and each local minima is assigned an exclusive tag. An intangible water drop is placed at each untagged pixel. The drop moves to low amplitude neighbor until it reaches a tagged pixel and it assumes tag value. In flooding approach, intangible pixel holes are pierced at each local minima. The water enters the holes and takings to fill each catchment basin. If the basin is about to overflow, a dam is built on its neighboring ridge line to the height of high altitude ridge point. These dam borders correspond to the watershed lines. Advantages of the watershed transform include the fact that it is a fast, simple and intuitive method. More importantly, it is able to produce a entire division of the image in separated regions even if the contrast is poor, thus there is no need to carry out any post processing work, such as contour joining. Its limitations will include over-segmentation and sensitivity to noise. There has also been an increasing interest in applying soft segmentation algorithms, where a pixel may be classified partially into multiple classes, for MR images segmentation. Clustering is a method of grouping a set of 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 attributes. Clustering technique is based on the computation of a measure of similarity or distance between the respective patterns. A cluster is a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters. Unlike classification, in which objects are assigned to predefined classes, clustering does not have any predefined classes. The main advantage of clustering is that interesting patterns and structures can be found directly from very large data sets with little or none of the background knowledge. The cluster results are subjective and implementation dependent. K-means

clustering is a simple clustering method with low computational complexity. The clusters produced by K-means clustering do not overlap. The Fuzzy C-means clustering algorithm is a soft segmentation method that has been used extensively for segmentation of MR brain images. In this work, we use K-means clustering and Fuzzy C-means Clustering methods exclusively to produce a primary segmentation of the image before we apply the marker controlled watershed segmentation algorithm.

K-means algorithm: K-means algorithm is under the category of Squared Error-Based Clustering (Vector Quantization) and it is also under the category of crisp clustering or hard clustering. K-means algorithm is very simple and can be easily implemented in solving many practical problems. K-Means is ideally suitable for biomedical image segmentation since the number of clusters (k) is usually known for images of particular regions of human anatomy. Steps of the K-means algorithm are given below:

- Choose k cluster centers to coincide with k randomly chosen patterns inside the hyper volume containing the pattern set.(C)
- Assign each pattern to the closest cluster center. ($C_i, i = 1, 2, \dots, C$)
- Recompute the cluster centers using the current cluster memberships.(U):

$$u_{ij} = \begin{cases} 1 & \text{if } \|x_j - c_i\|^2 \leq \|x_j - c_k\|^2, \text{ for each } k \neq i \\ 0 & \text{otherwise} \end{cases}$$

- If a convergence criterion is not met, go to step 2 with new cluster centers by the following equation, i.e., minimal decrease in squared error:

$$c_i = \frac{i}{|G_i|} \sum_{k, x_k \in G_i} X_k$$

where, $|G_i|$ is the size of G_i

or:

$$|G_i| = \sum_{j=1}^n u_{ij}$$

The performance of the K-means algorithm depends on the initial positions of the cluster centers. This is an inherently iterative algorithm. And also there is no guarantee about the convergence towards an

optimum solution. The convergence centroids vary with different initial points. It is also sensitive to noise and outliers. It is only based on numerical variables.

Fuzzy C-means algorithm: Fuzzy C-M clustering (FCM), also called as ISODATA, is a data clustering method in which each data point belongs to a cluster to a degree specified by a membership value. FCM is used in many applications like pattern recognition, classification, image segmentation. FCM divides a collection of n vectors c fuzzy groups and finds a cluster center in each group such that a cost function of dissimilarity measure is minimized. FCM uses fuzzy partitioning such that a given data point can belong to several groups with the degree of belongingness specified by membership values between 0 and 1. This algorithm is simply an iterated procedure. The algorithm is given below:

- Initialize the membership matrix U with random values between 0 and 1.
- Calculates c fuzzy cluster center $c_i, i = 1, \dots, c$, using the following equation:

$$C_i = \frac{\sum_{j=1}^n u_{ij}^m X_j}{\sum_{j=1}^n u_{ij}^m}$$

- Compute the cost by the following equation. Stop if either it is below a certain threshold value or its improvement over previous iteration:

$$J(U, C_1, \dots, C_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2$$

- Compute a new U by the equation. Go to step 2:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}}$$

There is no guarantee ensures that FCM converges to an optimum solution. The performance is based on the initial cluster centers. FCM also suffers from the presence of outliers and noise and it is difficult to identify the initial partitions.

Marker controlled watershed segmentation algorithm: Segmentation using the watershed transforms works well if you can identify, or mark, foreground objects and background locations. The gradient magnitude of the primary segmentation is

obtained by applying the Sobel operator. The Canny edge detector was also experimented on, but it was found that the results obtained by both methods are comparable. Hence, we decided on the Sobel filter as the Canny edge detector has higher complexity. In addition, the Sobel filter has the advantage of providing both a differencing and smoothing effect. Marker-controlled watershed segmentation follows this basic procedure:

- Compute a segmentation function. This is an image whose dark regions are the objects we are trying to segment
- Compute foreground markers. These are connected blobs of pixels within each of the objects
- Compute background markers. These are pixels that are not part of any object
- Modify the segmentation function so that it only has minima at the foreground and background marker locations.
- Compute the watershed transform of the modified segmentation function

MATERIALS AND METHODS

The proposed methodology is a two stage process. The first process uses K-means clustering to produce a primary segmentation of the input image, while the second process applies the marker controlled watershed segmentation algorithm to the primary segmentation to obtain the final segmentation map. Figure 1 describes the proposed method.

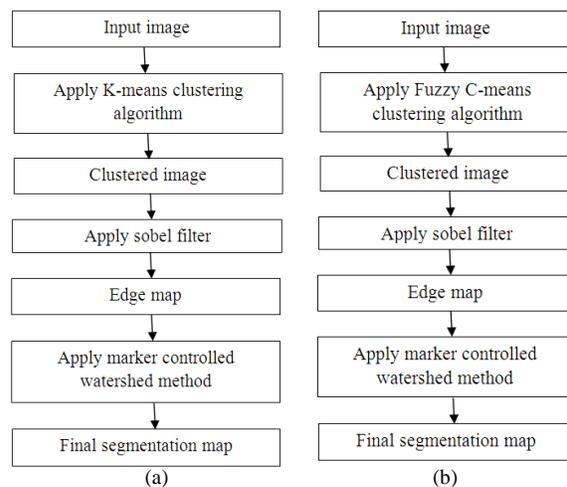


Fig. 1: Flow diagram of proposed method using (a) K-means clustering (b) Fuzzy C-means clustering

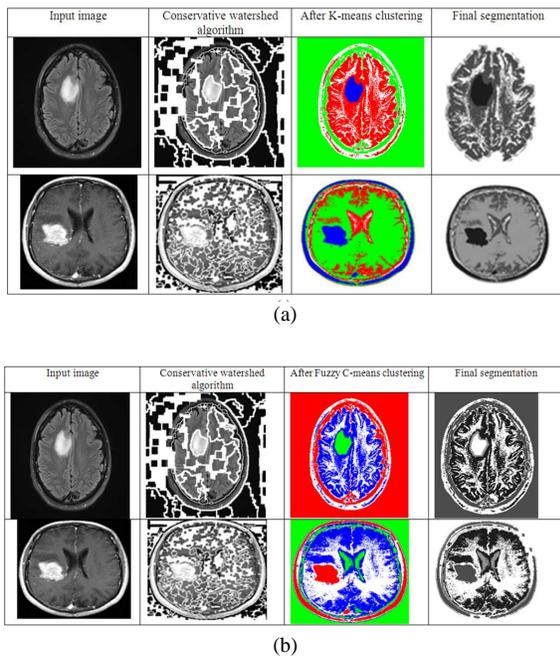


Fig. 2: Comparing segmentation maps obtained using (a) K-means (b) Fuzzy C-means clustering against conservative watershed method

RESULTS

We applied our proposed methodologies of K-Means clustering integrated with marker controlled watershed algorithm and Fuzzy C-means clustering integrated with marker controlled watershed algorithm to MR brain images of the head and obtained general segmentation maps of them. Figure 2 shows the segmentation results.

DISCUSSION

We evaluated the performance of our proposed methodologies by comparing them with conservative watershed algorithm. The use of K-means clustering and Fuzzy C-means clustering before applying marker controlled watershed segmentation algorithm had achieved the objective of reducing the problem of over-segmentation when applied to MR brain images.

CONCLUSION

Methods which integrated the K-Means clustering algorithm with the marker controlled watershed segmentation algorithm and Fuzzy C-Means clustering algorithm with the marker controlled watershed segmentation algorithm had been proposed. It

addressed the limitations of the conservative watershed algorithm, which included over segmentation. The experimental results had shown that our proposed methods of using K-means clustering and Fuzzy C-means clustering exclusively to obtain a primary segmentation of MR brain images before applying the marker controlled watershed segmentation to them is effective. And integration of K-means clustering with marker controlled watershed algorithm gave better segmentation than integration of Fuzzy C-means clustering with marker controlled watershed algorithm. By reducing the amount of over segmentation, we obtained a segmentation map which is more diplomats of the several anatomies in the medical images.

REFERENCES

- Christ, M.C.J. and R.M.S. Parvathi, 2011. An aadaptive mean-shift algorithm for MRI brain ssegmentation. *Int. J. Eng. Sci. Technol.*, 3: 5550-5554.
- Gonzalez, R.C., R.E. Woods and S.L. Eddins, 2004. *Digital Image Processing using MATLAB*. 1st Edn., Pearson Education India, India, ISBN: 9788177588989, pp: 620.
- Jayaraman, S., 2009. *Digital Image Processing*. 1st Edn., Tata McGraw Hill Education, New Delhi, ISBN: 9780070144798, pp: 723.
- Murugavalli, S. and V. Rajamani, 2007. An improved implementation of brain tumor detection using segmentation based on neuro fuzzy technique. *J. Comput. Sci.*, 3: 841-846. DOI: 10.3844/jcssp.2007.841.846
- Ng, H.P., S.H. Ong, K.W.C. Foong, P.S. Goh and W.L. Nowinski, 2006. Medical image segmentation using k-means clustering and improved watershed algorithm. *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI'06)*, IEEE Xplore Press, Denver, pp: 61-65. DOI: 10.1109/SSIAI.2006.1633722
- Shanmugam, N., A.B. Suryanarayana, D. Chandrashekar and C.N. Manjunath, 2011. A novel approach to medical image segmentation. *J. Comput. Sci.*, 7: 657-663. DOI: 10.3844/jcssp.2011.657.663
- Velmurugan, T. and T. Santhanam, 2010. Computational complexity between k-means and k-medoids clustering algorithms for normal and uniform distributions of data points. *J. Comput. Sci.*, 6: 363-368. DOI: 10.3844/jcssp.2010.363.368