

Automatic Segmentation of Brain CT Scan Image to Identify Hemorrhages

Bhavna Sharma

Department of Informatics and Computational Sciences, Mohan Lal Sukhadia University, Udaipur, India

K. Venugopalan

Professor, Department of Informatics and Computational Sciences, Mohan Lal Sukhadia University, Udaipur, India

ABSTRACT

Segmentation is required as a preliminary step in the analysis of medical images for computer aided diagnosis. For detecting tumors, edema, hemorrhage or any other abnormality given the complex structure of the brain, precise segmentation is crucial. CT scan is preferred method in traumatic brain injuries due to better contrast on bone, low cost and wide availability. This paper proposes fully automatic segmentation of brain CT images to identify hemorrhages. The method is comprised of three stages, preprocessing performed on the brain CT images, histogram based centroids initialization and finally the K-means clustering algorithm applied on the resultant image to segment the image in different clusters based on the intensity values of pixels. The method precisely segments the input image in exact clusters and analyzing those clusters hemorrhage can be identified.

Keywords

CT scan, Hemorrhage, Segmentation, Histogram, K-Means Clustering

1. INTRODUCTION

Segmentation refers to the process of partitioning an image into distinct regions by grouping together neighborhood pixels [1]. This is done based on specific properties or features of pixels representing objects in the image. Segmentation is an important step in image processing [2] and classification. Manual segmentation is also possible but it is a time consuming task and depends upon operator's variability. Therefore automatic methods are preferred over manual. In spite of several decades of research [3-4], automatic segmentation remains a challenging area in image processing. The demand for accurate segmentation methods has also grown stronger and consequently many segmentation methods have been developed based on different parameters of an image like gray level, depth, texture, color or motion [5].

Automatic segmentation of medical images is admittedly extremely challenging due to the inherent complexity of anatomical structures, deformation and non-rigid nature of organs. In addition, the image itself is dependent on the mode of image acquisition. Majority of medical images are gray scale representation, so segmentation is based on the gray level value of pixels. Medical image segmentation algorithms based on gray level values of pixels of an image can be divided into two categories: those based on similarities and those based on discontinuity. The former method involves segmenting an image based on similarity of intensity between pixels within a region while the latter uses sudden changes in gray level to indicate the discontinuity of a region [6].

Medical imaging techniques for brain scan like MRI, CT scan, PET scan are the tools used for extracting information by the radiologist. As compared to the other techniques used, CT scan is preferred because of wide availability, low cost and better contrast. In brain CT scan images for identification of hemorrhages [7] automatic and fast segmentation of image is a preliminary step [8] and thus choosing the accurate segmentation method becomes a crucial task.

2. METHODOLOGY

This study proposes a method for Brain segmentation for identifying and extracting hemorrhage region from 2D CT scan. The first step in the technique is preprocessing in which noise is removed from the image using median filtering or wiener filtering depending upon the type of noise present in the image. This leads to the second step which is the histogram based threshold calculation and finally applying K-Means clustering algorithm to divide the image into clusters. The whole procedure is fully automatic and the results are compared with the manual segmented regions done by the radiologist.

2.1 Preprocessing

Preprocessing is the initial step for detecting the anomaly in the brain CT scan image. This pre-processing step performs a non linear mapping of the gray level dynamics for the image. The use of filtering derives from the nature of the noise distribution in the CT images, such as white noise, salt and pepper type of noise. This can be removed by using Mean, Median filter or Wiener filters.

2.2 Automatic Thresholding

Clustering algorithms require an initialization of the cluster's centers. Usually, this is randomly made [9]. Nevertheless, an adequate selection allows for improvement in accuracy and also significantly reduces the number of required iterations to the convergence of these algorithms. The choice of the class number and initial correspondent centroids can be either supervised or unsupervised. The method of imposing a number and initial value of clusters in congruence with the quantity of information that we want to extract from the image is known as the supervised method. On the other hand the method that is based on the estimation of the number of clusters and initial cluster value in the image is termed as the unsupervised method. Among the methods used in this domain we consider the histogram information analysis. Histogram based thresholding is one of the most popular techniques for image segmentation because of its promising speed and simplicity. Peak detection algorithm also finds wide applications in image processing such as data deduction [10] threshold selection [11]. The algorithm determines peaks of image histograms and then automatically selects thresholds in

between those peaks. The algorithm consists of following steps [12]:

Step 1: Histogram definition for image size $S \times T$, at point (s, t) , $f(s, t)$ is the gray value with $0 \leq s \leq (S-1)$, $0 \leq t \leq (T-1)$. Let $H(g)$ denote the number of pixels having gray level g . Therefore, the histogram function can be written as:

$$H(g) = \sum_{s=0}^{S-1} \sum_{t=0}^{T-1} \delta(f(s, t) - g) \quad (1)$$

Where $g \in G$, $\delta(g=0) = 1$ and $\delta(g \neq 0) = 0$

- Step 2: Histogram smoothing to eliminate the parasitic peak.
- Step 3: Detect all local peaks. Local peak at position g satisfies the condition $H(g-1) < H(g)$ and $H(g) > H(g+1)$
- Step 4: Eliminating weak peaks among the detected peaks, there are some ones with weak heights; they represent small non significant regions. Eliminate all these peaks having minimal amplitude A_m .
- Step 5: After calculating the exact number of peaks, specifically finding the local and global minimum of peaks and the intensity values at which the peaks start, end, and attain their maxima are estimated. Gray-level thresholds are set between the two adjacent peaks.

Automatically calculated thresholds from this algorithm are the initial number of clusters, K for K-Means clustering algorithm. Correspondent's gray levels g_i of calculated thresholds are the initial centroids for those clusters, in K .

2.3 K-Means Clustering Method

It is one of the simplest unsupervised learning algorithms that solve the well known clustering problems [13]. In this procedure, a certain number of clusters are first assumed for a data set. For example, if K clusters are assumed, an equal number of centroid, one for each cluster is then defined. Each point of the data set is then associated to the nearest centroid and the entire set of centroids are then refined. The previous step is repeated for the new set of centroids thus setting up a loop. The entire steps are repeated until such point as no further changes result implying that the centroids do not move anymore and a stable result is obtained. Finally, this algorithm aims at minimizing an objective function, in this case a squared error function [14]. The objective function

$$J = \sum_{j=1}^K \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (2)$$

Where $\|x_i^{(j)} - c_j\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centers. The algorithm is composed of the following steps:

- Step 1: Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
- Step 2: Assign each object to the group that has the closest centroid.
- Step 3: When all objects have been assigned, recalculate the positions of the K centroids.
- Step 4: Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

3. EXPERIMENTAL RESULTS

The whole segmentation methods are implemented in MATLAB 7.1 on window XP and computer with 2.80 GHz CPU and 1 GB of RAM. The experimentation data consists of 30 brain CT images. Those images are passed as input to segmentation method and results are obtained based on the parameters. The selected images are of same quality but with hemorrhages of different sizes, shapes and at different locations.

In our image segmentation ground-truth value of images estimated by radiologist, is used as a reference against automatically segmented images, the performance is measured by calculating the discrepancy between the segmentation result and ground-truth segmentation. There are some key factors to judge the performance of an algorithm, Correctness is the ability to produce results that are consistent with ground truth, Stability is to achieve consistent results on different images given the same set of parameters and Computation time is the time taken by processor to complete segmentation of image.

This method gives correct segmentation of the hemorrhage region compared with the ground truth (manually segmented by radiologist). The original image is shown in Fig. 1(a) where as Fig. 1(b) is image after skull removal. Histogram of the image is shown in Fig. 2. The histogram with its peak values (Fig. 3) and calculated threshold values (Fig. 4) are also given. The segmented region of an epidural hemorrhage (biconvex shape with brightest intensity value of pixels) is shown in Fig. 5. This method is stable across images, given only the input image the method automatically calculates the initial thresholds. Number of threshold calculated are treated as the initial number of clusters for K-Means clustering algorithm and corresponding gray value of pixels are the initial centroids. The K Means algorithm calculates and divides the input image into three clusters shown in Fig. 6, where each cluster is having pixels value from input image and rest of pixels in the image is painted black. From all the clusters the brightest cluster can be extracted. That cluster can be further analyzed by considering its shape and size and the type of hemorrhage can be identified (in Fig.6 it is Epidural Hemorrhage).

The average computation time per image for segmenting the input image is 3sec.

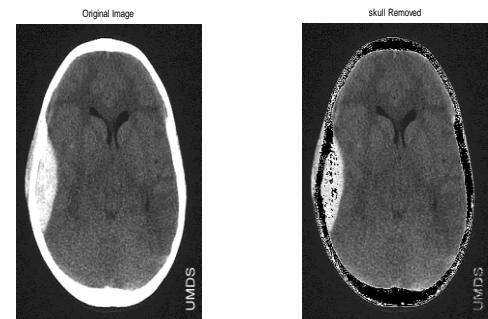


Fig. 1: (a) Original Image (b) Skull Removed

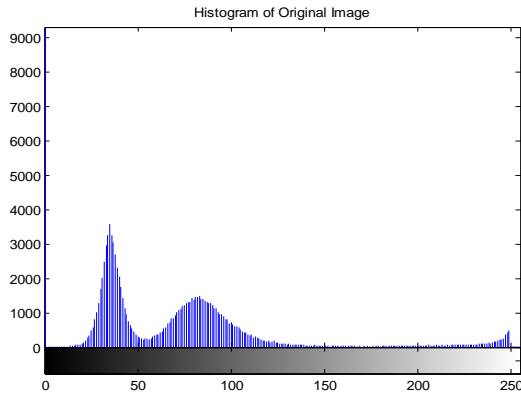
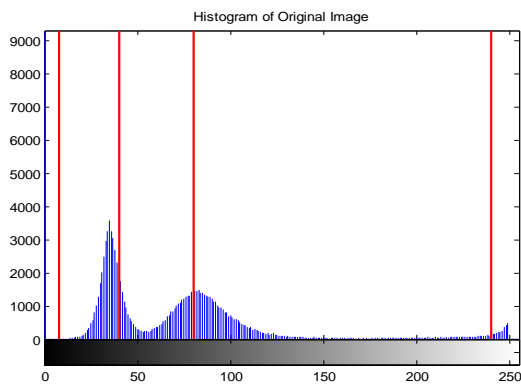
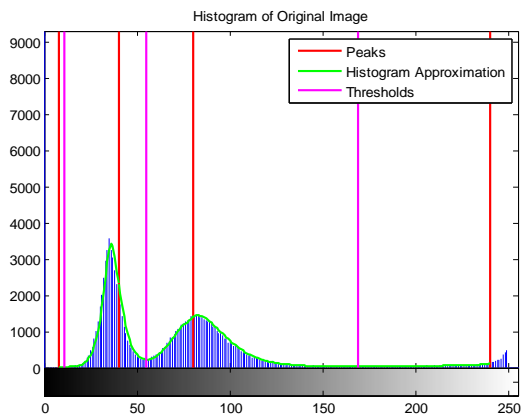


Fig. 2: Histogram of Original Image



The peaks of the histogram are at: 8 40 80 232

Fig. 3: Peaks of Histogram



The thresholds of the histogram are at: 11 55 170

Fig. 4: Thresholds of Histogram

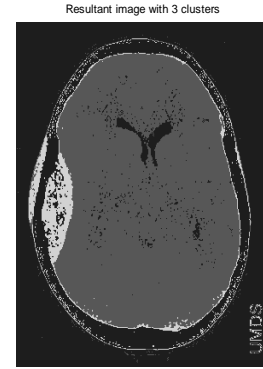


Fig. 5: Segmented Image with 3 clusters

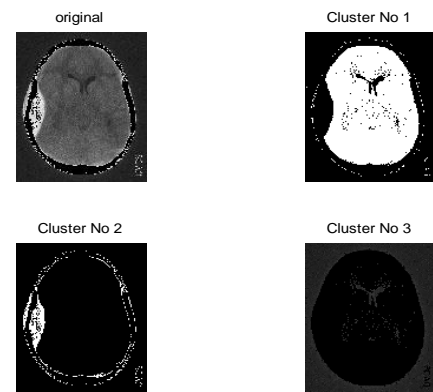


Fig. 6: Clusters of Original Image

4. CONCLUSION

This paper, presents a truly automatic image segmentation method. It can be regarded as truly automatic because it does not require a user to determine image-specific parameters such as thresholds or regions of interest. The use of the histogram analysis instead of a random initialization leads to an important improvement in the choice of the centers of clusters. The robustness of the method depends upon the different artifacts usually present in CT images such as noise and intensity in homogeneity. The future research in this field is to make this method more robust against all artifacts present in CT images and comparing this method with some well-known image segmentation methods.

5. ACKNOWLEDGMENT

We would like to thank Dr. Kapil Vyas, Radiologist, Geetanjali Medical College, Udaipur and Dr. Shailendra Pareek, Fortis Hospital, Jaipur, for their support lent to us during this study.

6. REFERENCES

- [1] Atam P. Dhawan, H. K. Huang, Dae Shik Kim , 2008. Principles and advanced methods in medical imaging and image analysis, World Scientific.
- [2] Rafael C. Gonzalez and Richard E. Woods, 2007. Digital Image Processing, 2nd ed., Beijing: Publishing House of Electronics Industry.
- [3] R.M. Haralick, L.G. Shapiro, "Image segmentation techniques", Computer Vision, Graphics and Image Processing, 1985, 29(1), 100-132.
- [4] N.R. Pal, S.K. Pal, "A review on image segmentation techniques", Pattern Recognition, 1993, 26(9), 1277-1294.
- [5] D.L. Pham, C. Xu, J.L. Prince, "Current methods in medical image segmentation", Annual Review of Biomedical Engineering, 2000, 2(1), 315-338.
- [6] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on System Man and Cybernetics, 1979, vol. 9, no. 1, 62-66.
- [7] W.L. Zhang, X.Z. Wang, "Extraction and classification for human brain CT images", Proceedings of Sixth International Conference on Machine learning and Cybernetics, IEEE 2007, 1155-1156.
- [8] S. M. Smith, "Fast robust automated brain extraction", Human Brain Mapping, 2002, vol. 17, 143-155.
- [9] Ruizhe Liu, Chew Lim Tan, Tze Yun Leong, Cheng Kiang Lee, Boon Chuan Pang, C. C. Tchoyoson Lim, Qi Tian, Suisheng Tang, Zhuo Zhang, "Hemorrhage slice detection in brain CT images", 978-1-4244-2175- 6/08 IEEE 2000.
- [10] M. I. Sezan, "A Peak Detection algorithm and its application to histogram-based image data reduction", Computer Vision, Graphics and Image Processing, 1990, vol. 49, 36- 51.
- [11] N. Papamarkos, B. Gatos, "A new approach for multilevel thresholds selection", CVGIP : Graphical Models and Image Processing ,1994,56(5), 357-370.
- [12] Bouchaib Cherradi, Omar Bouattane ,Mohamed Youssfi, Abdelhadi Raihani, "Brain Extraction and Fuzzy Tissue Segmentation in Cerebral 2D T1-Weighted Magnetic Resonance Images", International Journal of Computer Sciences Issues, 2011, vol. 8 issue 3.
- [13] S. Borah and M.K. Ghose, "Performance analysis of AIM-K-Means and K-Means in quality cluster generation", Journal of Computing, 2009, vol.1, issue 1, 2151-9617.
- [14] Alexander Rakhlin, Andrea Caponnetto, "Stability of K-Means Clustering", Advances in Neural Information Processing Systems 12, MIT Press, Cambridge, MA, 2007, 216–222.