International Journal of Computer Network and Security (IJCNS)
Vol 6. No.1 – Jan-March 2014 Pp. 69-73
©gopalax Journals, Singapore
available at : www.ijcns.com
ISSN: 0975-8283

HIERARCHICAL NORMALIZED CUT METHOD FOR CANCER IMAGE SEGMENTATION

Miss.Manochithra.P, Mrs.Latha Shunmugam.D,
Dr. S.J.S Paul Memorial College of Engineering and Technology,
Puducherry – 605 502
manochitra176@gmail.com, latha162@gmail.com

ABSTRACT

The image segmentation forms the basis for processing and analyzing an image required for most of the applications. This paper introduces an efficient method for detection of cancer in tissue. In the process of detection of cancer in tissue, segmentation plays vital role for partitioning an image into different subregion with homogeneous properties. The conventional methods use the hierarchical normalized cut (HNcuts) segmentation detects the cancer in tissue. HNcuts is a new minimally supervised hierarchical segmentation uses hierarchically represented data structure that merge two image segmentation algorithms – frequency weighted mean shift (FWMS) algorithm and normalized cut (Ncut) algorithm. HNCuts rapidly traverses a hierarchical pyramid, generated from the input image at various color resolutions, enabling the rapid analysis of images. HNcuts is robust, flexible, accurately quantifying the presence of cancer in tissue.

Key terms: Frequency mean shift (FWMS), normalized cut (Ncut), hierarchical normalized cut (HNcut)

1. INTRODUCTION

Segmentation is the very first step in almost all the image processing application where the properties of objects in image need to be analyzed e.g. in medical imaging problems in automotive vision in vehicle detection; object recognition in content based image retrieval etc. The objective of the image segmentation is to extract the dominant images. The image segmentation is very important to simplify an information extraction from images, such as color, texture, shape, and structure. The applications of image segmentation are diversely in many fields such as image compression, image retrieval, object detection, image enhancement, and medical image processing. Image segmentation technique applied to tissue for detecting cancer. There are many techniques to segment an image into regions that are homogenous. As the structure of medical image is inaccurate and complex these techniques are not suitable for their analysis in order to extract the useful features. Clustering plays crucial role while executing the task of organizing the objects into groups based n its properties. A cluster is therefore a set of objects which are analogous between them and are dissimilar to the objects belonging to other clusters [8]. The concept of mean shift estimating the probability density gradient function is discussed by Cheng Yizong [2] according to which mean shift algorithm has been extensively applicable in the area of computer vision like tracking, image segmentation edge extraction, motion estimation, feature space analysis, filtering, video analysis and for other numerous tasks, on the other hand this algorithm cause failure because of its over segmentation phenomenon. Graph cut segmentation is the another development for solving the crucial task of image segmentation in which pixels are considered as nodes and unidirected weighted graph is used while segmenting an image by using Graph cut method. While partitioning the graph, deciding precise criterion for good partition and computation of such partition are the primary requirements. In order to resolve these requirements image segmentation approach based on Normalized cut (Ncut) has been proposed by Shi and Malik [1]. Ncut algorithm finds wide acceptance not only in the field of image processing but also in related fields like motion picture, medical imaging and vector field of segmentation. However this algorithm fails when there are more pixels in the image due to which number of graph nodes are generated which causes complexity to solve the algorithm. To overcome the normalized cut (Ncut) disadvantage, mean shift and normalized cut are combined together [5]. The

International Journal of Computer Network and Security (IJCNS)
Vol 6. No.1 – Jan-March 2014 Pp. 69-73
©gopalax Journals, Singapore
available at : www.ijcns.com
ISSN: 0975-8283

frequency weighted mean shift (FWMS) [6] which improves the computational time over that of mean shift, so we combine FWMS and Ncut which improves the time of segmentation.

The rest of paper is organized as follows. In section II we describe the segmentation technique of our work. In section III, we describe the methodological description of HNcut. Finally concluding the remarks are presented in section IV.

2. SEGMENTATION TECHNIQUE

For the extraction of useful features from the complex tissue, Tissue microarray (TMA) analysis is reliable. TMA is very important in order to improve the diagnosis and treatment of brain tumor, by detecting cancer in tissue at its early stage. Segmentation of medical images is first important step in their analysis, the segmentation gives organ detection and variation of growth of tissues as a output in medical images. In this paper hierarchical normalized cut segmentation techniques are discussed as below.

Frequency Weighted Mean Shift

FWMS algorithm is based on mean shift algorithm [2]. By using ms algorithm we detect the modes using density gradient. We start with the fixed point iteration update ∀j ∈ {1. . . N} in MS algorithm

$$f_{k+1} = \frac{\sum_{i=1}^{N} f_{k,i}\ G(f_{k,j} - f_{k,i})}{\sum_{i=1}^{N} f_{k,i}\ G(f_{k,j} - f_{k,i})} \qquad (1)$$

Where G is a Gaussian function with a bandwidth parameter σ_{MS} which is used to compute the kernel density estimate at data point. We can furthermore reduce the computational cost in ms algorithm O(N^2) to O(N) by employing the IFGT[4], we can reduce the computation complexity.

It becomes possible to exploit the fact that after each iteration of the MS many of the data points, in our case color value converge. If we consider what the convergence means mathematically, essentially two points c_{β_1}, c_{β_2} where $\beta_1, \beta_2 \in \{1, \ldots, N\}$ meet the requirement $\left|f_{k,\beta_1} - f_{k,\beta_2}\right|$ that where is a predefined tolerance value. We can just rewrite the numerator of (12), which is

$$f_{k,\beta_1} G(f_{k,j} - f_{k,\beta_1}) + f_{k,\beta_2} G(f_{k,j} - f_{k,\beta_2})$$

$$+ \sum_{i=1}^{N} f_{k,i}\ G(f_{k,j} - f_{k,i}) \qquad (2)$$

$$2 f_{k,\beta_1} G(f_{k,i} - f_{k,\beta_1}) + \sum_{i=1}^{N} f_{k,i}\ G(f_{k,j} - f_{k,i}) \quad (3)$$

There by avoiding the explicit calculation of $G(f_{k,j} -$ fk,β2, where j, β1, β2 ∈ {1. . . N}, k ∈ {1. . . K}. This results in one less computation for the Gaussian, which is by far the most expensive operation in the entire MS clustering process. The formulation in (3) results in a significant computational efficiency improvement. The computational savings apply to the denominator as well, as it follows the same reduction.

As a result, we may rewrite the update presented in (1) as a multistep update. Initially, we determine the unique values in F_K under the constraint that any color values $\left|f_{k,i} - f_{k,j}\right|$ are considered equivalent. Thus, from $F_K = \{f_{k,1}, f_{k,2}, f_{k,3} \ldots |F_K|\}$ we can construct the vector $\hat{F}_k$, where $\hat{F}_k \subset F_k$ and $\hat{F}_k$ is a set of only unique values in F_K, with $\left|\hat{F}_k\right| = M_k$. A weight vector $w_K = \{w_K, 1, \ldots, w_K, M_k\}$ is then computed for $\hat{F}_k$ as

$$w_{k,j} = \sum_{i=1}^{|F_k|} w_{k-1,i} \qquad (4)$$

Where j ∈ {1. . .M_k }. Equation (4) is summing the weights from the previous level into the new unique values that resulted from the next iteration of mean shifting. As a result, $w_{k,j}$ contains a count of the number of original pixels that have migrated to $F_{k,j}$ through mean shifting. When k = 1, we define w_0 as a vector of length *N*, filled with ones, representing that each color value has equal weighting. Now, the number of points in the system that have converged to some intensity (color) value $\hat{f}_{k,j}$ is represented by $w_{k,j}$. It is important to note the following definition of M_k where

$$|w_k| = \left|\hat{F}_k\right| = |F_{k+1}| \qquad (5)$$

$$\sum_{i=1}^{M_k} w_{k,i} = N \qquad (6)$$

This leads us to the update of

$$f_{k+1} = \frac{\sum_{i=1}^{M_k} w_{k,i}\hat{f}_{k,i}\, G(f_{k,j} - f_{k,i})}{\sum_{i=1}^{M_k} w_{k,i} G(f_{k,j} - f_{k,i})} \qquad (7)$$

Normalized Cut

International Journal of Computer Network and Security (IJCNS)
Vol 6. No.1 – Jan-March 2014 Pp. 69-73
©gopalax Journals, Singapore
available at : www.ijcns.com
ISSN: 0975-8283

Normalized cut [1] is a global criterion for partitioning the graph used for segmentation of image. Normalized cut criterion measures both the total dissimilarity and similarity between different groups. A graph *G = (V, E)* can be partitioned into two disjoint sets, *A*, *B*, *A U B*, *=V*, *A* intersection *B*= *Ø* by simply removing edges connecting the two parts. The degree of dissimilarity been removed. In graphical language, it is called the cut.

$$cut(A,B) = \sum_{\mu \epsilon A, v \epsilon b} w(u,v) \qquad (8)$$

The successful bipartitioning of a graph is the done when it minimizes this cut value. Although there are a various number of such partitions, in the past lot of work was done for finding the minimum cut of a graph. Wu and Leahy [9] proposed a clustering method based on this minimum cut criterion. They partition a graph into k-sub graphs such that the maximum cut across the subgroups is minimized. By finding the minimum cut this problem can be efficiently solved by them. However the minimum cut criteria favors cutting small sets of individual nodes in the graph. Fig. 1 illustrates one such case

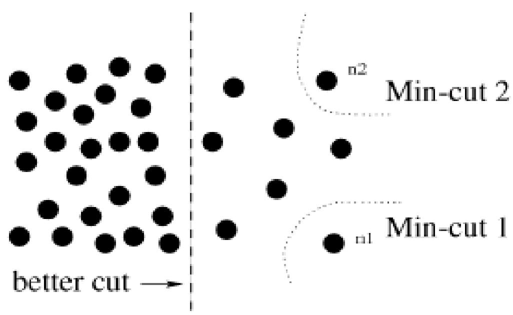


Fig ure.2.1 An example illustrating that minimum cut gives a bad partition.

From figure 3.1 any cut those partitions out individual nodes on the right half will definitely have smaller cut value than the cut that partitions the nodes into the left and right halves. To avoid this problem of partitioning small sets of points, [1] propose a new measure of finding cut between two groups. Instead of looking at the value of total edge weight connecting the two partitions, they computes the cut cost as a fraction of the total edge connections to all the nodes in the graph and call this disassociation measure the normalized cut (Ncut).

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \qquad (9)$$

Where *Assoc (A,V)* = $\sum_{\mu \in A, t \in B} w(u,t)$ is a total connection from nodes A to all nodes in the graph and *assoc(B,V)* is similarly defined. However if there are more pixels in the image, more graph of nodes will be generated and this will cause difficulties to solve this algorithm.

3. PROPOSED APPROACH

When we apply graph partitioning techniques directly on images, it has appeared to be inefficient and more complex with low operating speed. If Ncut is directly applied on image pixels, more graph nodes will be generated and this will cause difficulties to perform the segmentation. To overcome the limitations of Normalized cut (Ncut) algorithm by using the merits of Frequency weighted mean shift algorithm, integration of both Frequency weighted mean shift and Normalized cut algorithms is done in the proposed approach.

We start by requiring the user to select a few sample pixels from the target class from an image. We use these pixels to guide the subsequent pixel classification process across all images in the same domain. Next, we employ the MS algorithm on the color values in the image to form a hierarchical data structure. Intuitively, the FWMS algorithm allows for identification of color values which are within some specified tolerance of each other and assigns them to the same mode. Employing the NCut to FWMS output only on to the unique color values which results only fewer computations. The above algorithm is implemented in cancer in tissue.

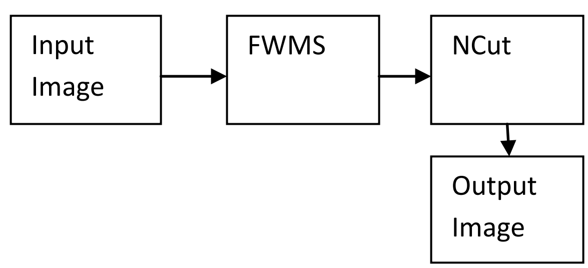


Figure 3.1: Block Diagram of Proposed Algorithm

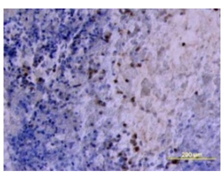
4. EXPERIMENTAL RESULTS AND DISCUSSIONS

When Normalized cut (Ncut) method is directly applied on images, it takes more time for segmenting an image because of more number of graph

International Journal of Computer Network and Security (IJCNS)
Vol 6. No.1 – Jan-March 2014 Pp. 69-73
©gopalax Journals, Singapore
available at : www.ijcns.com
ISSN: 0975-8283

nodes are generated . This problem can be overcomed by using frequency weighted mean shift algorithm before applying an Ncut algorithm to get results in less time. In this work we provide a improved image segmentation method for determining the cancer in tissues which calculate the peak signal to noise ratio.

Experiments were performed by taking a variety of images from different planes and proposed image segmentation algorithm was applied on it. First experiment was performed by taking input image of 156 x 156 pixels, by using the proposed segmentation algorithm; we detect cancer in tissue more accurately with PSNR value of 7.48 and time taken to segment is 0.6s.



(a)

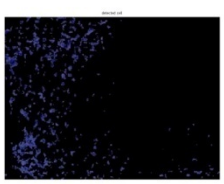


Figure 4.1: (a) input image (b) detection of cancer by HNCut

5. CONCLUSION

Frequency weighted mean shift algorithm is an efficient method of clustering, which segments the image into multiple separated color value regions having homogeneous properties. However on applying Ncut method on images it takes more time for segmentation because of more graph nodes are generated which causes the difficulties to solve this algorithm. In this paper an image segmentation algorithm has been implemented and it is based on the conventional Frequency weighted mean shift algorithm and Ncut algorithm. The effectiveness of the proposed algorithm have verified by some experimental results to express its improved performance in detecting the cancer and calculating the parameters of the image. In future we can combine improved watershed algorithm and normalized cut algorithm for detecting cancer. This image segmentation technique will be accurate and efficient for detecting cancer and also overcomes the drawbacks of normalized cut.

REFERENCES

[1]. Shi J and Malik J, “Normalized cuts and image segmentation”, IEEE Trans, on PAMI, Vol. 22, no. 8, pp. 888-905, 2000.

[2]. Cheng Y, “Mean shift, mode seeking, and clustering”, IEEE Trans, on Pattern Analysis and Machine Intelligence, Vol. 17, no. 8, pp. 790-799, 1995.

[3]. Qiu-Bo Xi, “An Improved Image Segmentation Algorithm Base on Normalized Cut”, 2nd International Conference on Computer Eng. & Tech., vol (7), pp. 293-296, 2010.

[4]. C. Yang, R. Duraiswami, N.A. Gumerov, and L. Davis, “Improved fast Gauss transform and efficient kernel density estimation”. in Proc. IEEE Int. Conf. Comput. Vis., Oct. 2003, vol. 1, pp. 664–671.

[5]. W. Tao, H. Jin, and Y. Zhang, “Color image segmentation based on mean shift and normalized cuts,” IEEE Trans. Syst., Man, Cybern., vol. 37, no. 5, pp. 1382–1389, Oct. 2007.

[6]. Janowczyk, S. Chandran, R. Singh, D. Sasaroli, G. Coukos, M. D. Feldman, and A. Madabhushi, “Hierarchical normalized cuts: Unsupervised segmentation of vascular biomarkers from ovarian cancer tissue microarrays,” in Proc. Med. Image Comput. Assist. Intervention, 2009, vol. 1, pp. 230–238.

[7]. Comannicau D. “Mean Shift: A Robust Approach toward Feature Space Analysis”, IEEE Transactions on pattern analysis and machine intelligence, Vol.24, No. 5 May 2002.

[8]. Thilagamani S.and Shanthi N, “A Survey on image segmentation through clustering”, International journal of research and Information sciences vol 1 No. 1, March 2011.

[9]. Wu Z,.Leahy R, "An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation", IEEE Transaction on pattern analysis and machine intelligence Vol. 15, NO. 11, pp. 1101 – 1113, Nov 1993.

[10]. X. Yuan, B. G. Hu, and R. He, "Agglomerative mean-shift clustering via query set compression," in Proc. SIAM Int. Conf. Data Mining, 2009, pp. 221–232.