International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org || PP. 27-32

Hierarchical Numerical Normalized Cuts Algorithm For Detecting Tissue Microarrays

Dr. C.Bhuvaneswari[1], S.Daisy Fathima Mary[2]
[1]Assistant professor & Head, [2]Lecturer
1,2 Department of computer science , Thiruvalluvar University model college of arts and science
Thiruvennainallur, Tamilnadu, India
bhuvana.csdept@gmail.com, daisyfatima_mca@yahoo.co.in

Abstract: *The identification of the extent and intensity of the stain could provide a quantitative and reproducible prognostic metric that could help predict risk of ovarian cancer recurrence and patient survival. The HNCut results be deemed to have acceptable accuracy. The Hierarchical Normalized Cuts algorithm identifies the cancer tissues of the derived input image from the CTScan. The Frequency Weighted Mean Shift algorithm identifies the color values which is decomposed into multiple levels of color resolution. The NCuts is a graph partitioning method is used to separate data into disjoint sets and gets the cancer tissue pixel as gray scale image and eliminates the remaining pixels. The combination of FWMS and NCuts is termed as HNCuts is precise and accurate quantification of extent of vascular staining of ovarian cancer. The Hierarchical Normalized Cut method is employed and output is achieved for the same with the computation time is calculated.*

Keywords: *CTScan, Frequency Weighted Mean Shift (FWMS)algorithm, HNCuts, NCuts, segmentation.*

I. Introduction

A method is applied to analyze gene expression in digital images of TMA data from tissue sections processed for detection of specific gene transcripts. The goal of this approach is to enable identification of local regions of differential expression and to provide a means of ranking these regions across anatomic and cellular substrates. Most previous computerized image analysis algorithms for TMAs have involved thresholding-based schemes. These methods are known to be highly sensitive to even slight changes in color and illumination. Clustering-based approaches, including k-means, have also been investigated for the analysis of TMAs. However, k-means is a nondeterministic algorithm and is highly sensitive to the initial choice of cluster centers. Active contour schemes, while suitable for cell and nuclear segmentation in digital pathology, are not ideally suited to the problem of pixel-level classification. Additionally, they are typically infeasibility where hundreds of objects need to be concurrently segmented on very large images. The present approach segments and isolates gene expression signal through methods associated with image analysis, pattern recognition. Digital Imaging and Communications in Medicine (DICOM) is a standard for handling, storing, printing, and transmitting information in medical imaging. It includes a file format definition and a network communications protocol. DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format. DICOM enables the integration of scanners, servers, workstations, printers, and network hardware from multiple manufacturers into a picture archiving and communication system (PACS).

II. Literature Review

"The HNCut algorithm is more flexible compared to supervised schemes in its ability to segment different object classes. The combination of both the high-throughput efficiency and flexibility of HNCut makes it ideally suited to applications requiring high-throughput analysis, such as quantifying the expression of biomarkers on TMAs. In this paper, we demonstrate the specific application of HNCut to a problem of automated quantification of stain extent associated with a vascular marker on OCa TMAs. The mean-shift algorithm (MS) has been employed and modified as an unsupervised technique for mode discovery instead of k-means. The MS algorithm attempts to identify the cluster mean within a predefined bandwidth"[1].

"Angiogenesis, the formation of blood vessels, is a process whereby capillary sprout are formed in response to external stimuli. We model the tumor induced angiogenesis on keys events such of migratory response of endothelial cells to tumor angiogenic factors and the local cell interaction with the extracellular matrix (ECM). We consider the ECM medium as a statistically inhomogeneous two-phase random medium. Numerical simulations of the model are presented. Using this model, we will compare the influence of ECM distribution on vascular network formation. By developing mathematical models of angiogenesis, we hope to provide a deeper insight into the mechanisms underlying angiogenesis"[2].

“Automated detection and segmentation of nuclear and glandular structures is critical for classification and grading of prostate and breast cancer histopathology. In this paper, we present a methodology for automated detection and segmentation of structures of interest in digitized histopathology images. The scheme integrates image information from across three different scales: (1) low- level information based on pixel values, (2) high-level information based on relationships between pixels for object detection, and (3) domain-specific information based on relationships between histological structures. Low-level information is utilized by a Bayesian classifier to generate a likelihood that each pixel belongs to an object of interest [3]”.

“Quantifying expression levels of proteins with sub cellular resolution is critical to many applications ranging from biomarker discovery to treatment planning. In this paper, we present a fully automated method and a new metric that quantifies the expression of target proteins in immunohisto-chemically stained tissue microarray (TMA) samples. The proposed metric is superior to existing intensity or ratio-based methods. We compared performance with the majority decision of a group of 19 observers scoring estrogen receptor (ER) status, achieving a detection rate of 96% with 90% specificity. The presented methods will accelerate the processes of biomarker discovery and transitioning of biomarkers from research bench to clinical utility.[4]”.

III. Tissue Microarrays

The TMA technology is to address the limitations of conventional techniques and to enable ‘genome-scale’ molecular pathology studies. TMAs facilitate the analysis of molecular alterations in thousands of tissue specimens in a massively parallel fashion. Construction of TMAs is achieved by acquiring cylindrical core specimens from up to 1000 fixed and paraffin-embedded tissue specimens and arraying them at high density into a recipient TMA block. Up to 300 consecutive sections can be cut from each TMA block and probed with detection reagents for a variety of molecular targets either at the DNA, RNA or protein level. A single TMA experiment can yield information on the molecular characteristics of up to 1000 specimens at once. This is in contrast to conventional analyses, where each slide contains a section of a single tissue.

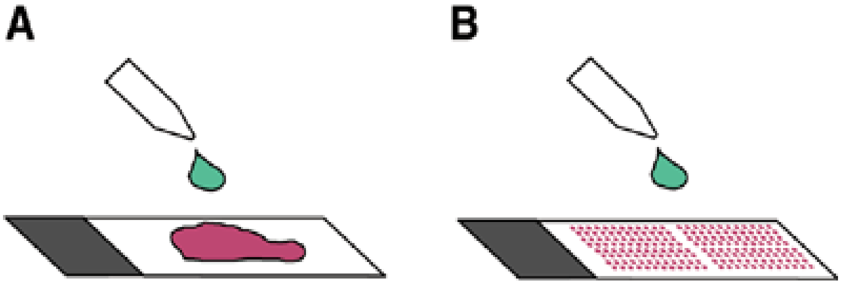


Fig. 1. Comparison of conventional section analysis.

. To identify TVMs, genes that fulfilled at least one of the following three criteria were identified: the gene was present in at least 15 of 21 TVC samples and absent in at least three of four normal vascular samples; or was present in at least 18 of 21 TVC samples and was at least three-fold over expressed in TVC relative to normal vascular samples; or was expressed at the highest values in the TVC samples (ie, at least 20 of 21 TVC samples had higher expression values relative to normal samples). Seventy genes emerged for this analysis .reflections, and masking portions of images.

IV. Methodology

Hierarchical Normalized Cuts

The algorithm is specifically designed for rapid extraction of pixels of interest in a minimally supervised manner as opposed to unsupervised clustering which is insensitive to the user’s domain knowledge as the aforementioned approaches take. Input images are in the form of DICOM images. These DICOM images are converted into the JPEG images are considered as input image. The desired target class based on individual representative colors selected from the target class by a user. Then a few selected sample pixels from the target class from an image are used to guide the subsequent pixel classification process across all images in the same domain.

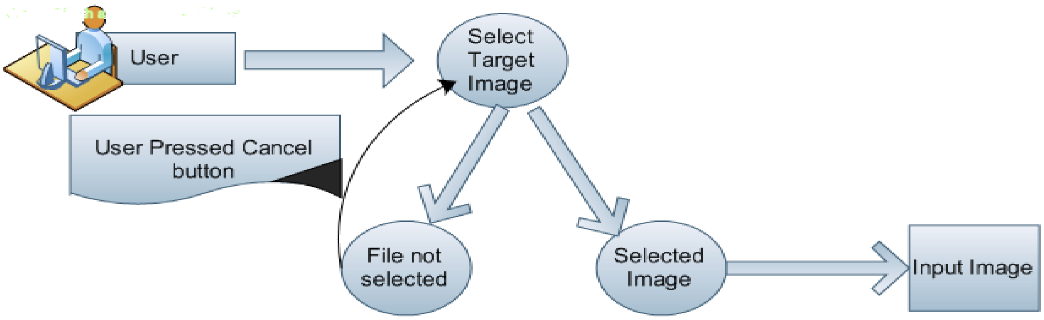


Fig. 2. Input Image Block Diagram.

Frequency Weighter Mean Shift Algorithm

It represents itself as randomly occurring white and black 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 NCuts Operation only on the unique values at each level, all possible color values, allows for a factorization resulting in significantly fewer computations. An illustration that computes the weight for each unique mode which reflects the actual frequency of the number of pixels associated with it. The MS algorithm is used to detect modes in data using a density gradient estimation. By solving for when the density gradient is zero and the Hessian is negative semi definite we can identify local maxima.

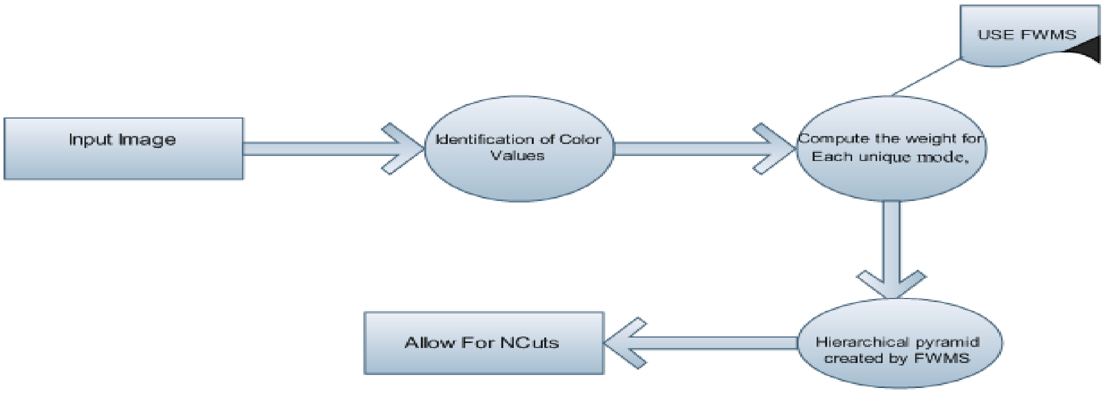


Fig. 3 Frequency Weighted Mean Shift Algorithm Block Diagram

Algorithm:

1. 1.The first is the maximum number of iterations which is specified by the user.
2. The second more common approach is to stop the algorithm when the difference between any two iterations falls below a predefined threshold (i.e., the amplitude of the migrations associated with each point reduces significantly).
3. Finally, convergence can be reached when the number of elements in ˆF becomes small enough that additional clustering provides no efficiency benefit as the overhead in the NCut starts to outweigh the computation time.

Normalized Cuts

Normalized Cuts (NCuts) is among the final mature Descendants from a series of graph-cutting techniques ranging from maxcut to mincut. NCuts is a graph partitioning method used to separate data into disjoint sets. In our cases, the hierarchical pyramid created by FWMS at various levels of color resolution (ˆF1 ,ˆF2, . . . ,ˆFK) where K represents levels of the color pyramid serves as the initial input to the Ncut algorithm. The NCut takes a connected graph G= (E, V), with vertices (V) and edges (E) and partitions the vertices into disjoint groups. By setting V equal to the set of color values ˆFK and having the edges represent the similarity between the color values. The NCut is defined as the process by which the removal of edges leads to two disjointed.

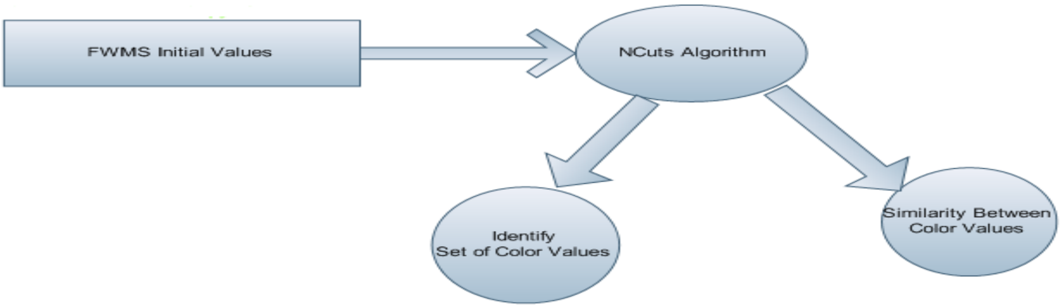


Fig. 4 Block Diagram of Normalized cu

The application of NCut on the color values that have not been eliminated are mapped back to the spatial domain via their original pixel locations.

Hierarchcal Numerical Normalized Cuts

A new minimally supervised hierarchical segmentation approach that combines a FWMS and normalized cuts (HNCuts) for pixel-level detection and classification. HNCut is able to segment very large images rapidly. HNCut is largely insensitive to choice of parameter value and is able to discriminate between regions with similar color values. The parameters for NCuts are automatically computed, and the parameters for the FWMS are automatically adjusted based on the variance of the output.

Segmented Image

Segmentation scheme termed HNCuts for precise and accurate quantification of extent of vascular staining of OCa TMAs. The extent and severity of this vascular stain has been predicted to be an important prognostic marker in predicting outcome of women with OCa. The strength of HNCut is derived from the fact that it integrates the best of both an FWMS clustering and the NCut algorithm. While other schemes have been previously proposed in an attempt to combine both mean shift and NCuts, we believe that HNCut is the only approach which provides the flexibility to be able to extract different target classes based on user input.

V. Experiment And Results

The cancer tissues in an ovarian by using a Hierarchical Normalized cuts methods with the combination of two methods are Frequency weighted Mean Shift methods and Normalized cuts methods are sed to identify the cancer tissues.

A. Query Image

The first step is getting an input image by browsing the data sets. This image is an unmarked cancer image.Figure captions appear below the figure, are flush left, and are in lower case letters. When referring to a figure in the body of the text, the abbreviation "Fig." is used. Figures should be numbered in the order they appear in the text.

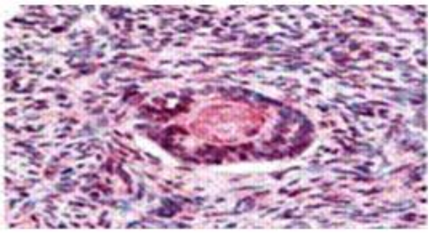


Fig 5. Input image

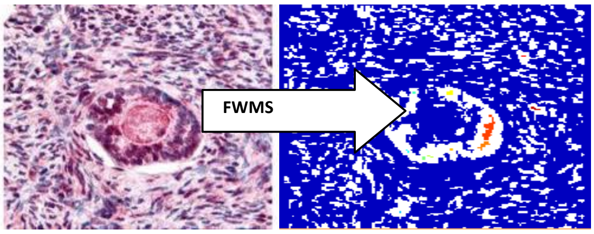
The image database consist of 60 ovarian cancer tissues images from 15 patients. These images are taken from CT Scan and converted into JPEG images.

B. Identifying The Cancer Tissues

The two steps that are involved to identify the cancer tissues are Frequency Weighted Mean Shift and Normalized cut.

a. Frequency Weighted Mean Shift

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 NCuts Operation only on the unique values at each level, all possible color values, allows for a factorization resulting in significantly fewer computations.



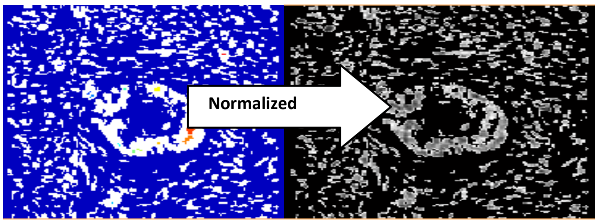
INPUT IMAGE **FWMS IMAGE**

Fig 6. Preprocessing Model

After applying this algorithm, the input image is decomposed into multiple levels of color resolution, which is the result of this phase.

b. Normalized cut

NCuts is a graph partitioning method, used to separate data into disjoint sets. Normalized Cuts (NCuts) is among the final mature Descendants from a series of graph-cutting techniques ranging from maxcut to mincut .

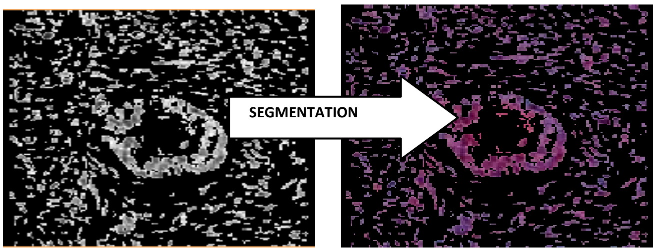


FWMS IMAGE **NORMALIZED IMAGE**

The application of NCut on the color values that have not been eliminated are mapped back to the spatial domain via their original pixel locations. After applying the NCuts methods we get the cancer tissue pixel as gray scale image and eliminating the remaining pixels. The combination of FWMS and NCuts is termed as HNCuts

C. Segmented Output

Segmentation scheme termed HNCuts for precise and accurate quantification of extent of vascular staining of OCa TMAs. The extent and severity of this vascular stain has been predicted with OCa.



NORMALIZED IMAGE **SEGMENTEDIMAGE WITH OC TISSUES**

The final stage of this thesis is the segmented output. Here the cancer tissue is predicted accurately from the input image.

D. Computation Process

The computation value is calculated for the work from the first step to final step. According to the input image, the computation value varies. The computation value is calculated by computing the edges of the images based on intervening contours, NCuts eigenvector and discrete segmentation. By choosing above image as input image, the computation value calculated is 0.0036745 in seconds. The cancer tissues are identified and highlighted by using HNcut methods from the input image and the computation value is also calculated for it.

VI. Conclusion

A minimally supervised segmentation scheme termed as Hierarchical Normalized Cut for precise and accurate quantification of extent of vascular staining of ovarian cancer. The extent and severity of this vascular stain has been predicted to be an important for prognostic marker. The strength of Hierarchical Normalized Cut is derived from the fact that it integrates the best of both a Frequency Weighted Mean Shift and the Normalized Cut algorithm. Hierarchical Normalized Cut is the only approach which provides the flexibility to be able to extract different target classes based on user input. The 60 data set of ovarian cancer tissues taken as input and Hierarchical Normalized Cut method is employed and output is achieved for the same with the computation time is calculated. In future work it intends to explore the applicability of Hierarchical Normalized Cut to other biomarker quantification and digital pathology problems. And also Hierarchical Normalized Cut operates in the color space, and is thus highly efficient, no limitation in the size of the image that can be analyzed by Hierarchical Normalized Cut is the amount of computer memory available to read in the image data.

References

[1]. High-Throughput Biomarker Segmentation on Ovarian Cancer Tissue Microarrays via Hierarchical Normalized Cuts; Andrew Janowczyk, Sharat Chandran, Rajendra Singh; IEEE Trans. Biomed. Eng., vol.59, no. 5, May.2010.

[2]. Stochastic modeling of tumor induced angiogenesis in a heterogeneous medium, the extracellular matrix; A. Rabinovich, S. Krajewski; Eng. in Med. and Bio. Soc., pp. 3146 – 3149, Sep. 2006.

[3]. Automated Gland and Nuclei Segmentation For Grading Of Prostate And Breast Cancer Histopathology; M. Krajewska, A. Shabaik; IEEE Trans. Biomed., pp. 284-287,May. 2008.

[4]. Quantification of Sub cellular Molecules in Tissue Microarray; J. Wu, J. Dong, and H. Zhou; Proc. Soc. Photo-Opt. Instrum. Eng., Mar. 2006,vol. 6143, pp. 509–520.

[5]. Expectation Maximization driven Geodesic Active Contour with Overlap Resolution (EMaGACOR): Application to Lymphocyte Segmentation on Breast Cancer Histopathology; A. Rabinovich, S. Krajewski; IEEE Trans. Biomed. Eng., vol.57, no. 7, pp. 1676 - 1689, Jul.2010.

[6]. Predicting Classifier Performance With A Small Training Set: Applications To Computer-Aided Diagnosis And Prognosis; Y. Freund and R. Schapire; IEEE Trans. Biomed., pp. 229-232,Apr. 2010.