

Segmentation of Lung Nodule Image Using Global Thresholding and Classification by K-NN

Isaac jeba Kumar¹, Selvaganesh²

P.G. Student, Sethu Institute of Technology, Kariapatti, Tamil Nadu, India¹

Associate Professor, Sethu Institute of Technology, Kariapatti, Tamil Nadu, India²

ABSTRACT: We present a new frame of lung parenchyma Segmentation. The parenchyma execution is based on global threshold value method and boundary tracking method. The parenchyma scheme to represent the lower CT value of overall lung. The parenchymal lung images segmentation was based primarily on that the CT lung parenchyma density was lower, while the pleura surrounding soft tissue, mediastinum, bone and others' are higher. The automatic segmentation process in this article includes global threshold binarization, the boundary is extracted to remove the background of the trunk area, threshold binarization after get rid of the background, lung area repair lung parenchyma extraction, and so on. The optimal threshold value method for each site CT images to automatically generate optimal threshold value. After threshold segmentation, there is still strip external interference in the outside region of the human trunk. To perform segmentation of the lungs using a technique that selects an optimal threshold for a given cancerous patient by comparing the curvature of the lung boundary to that of the ribs. Optimal thresholding to automatically select a segmentation threshold for the image volume. The segmented CT image is classified using the K-NN classifier

KEYWORDS: CT images, Thresholding, Global thresholding, KNN

I. INTRODUCTION

One of the basic and most representational algorithms in image segmentation is Thresholding segmentation. For the greyscale image in which the objects and background are in varied level, the specified segmentation is intuitionistic and easily applicable [1]. It is divided into two types global threshold and local threshold. If the pixels in the image are segmented with same threshold, then it is known as global threshold method; if the pixels are segmented with different thresholds, it is known as local threshold method. Choosing the threshold value is the main key of thresholding segmentation. Threshold can be chosen easily by making it a constant. In order to obtain a fixed, sufficient and better segmentation, the background and foreground of image should be in various gray levels [2]. One of the most popular in image segmentation techniques is multilevel thresholding. It is a straightforward and effective approach. It is also used in the field of Pattern recognition and computer vision. Multilevel Thresholding is divided into two main groups: Parametric and non-parametric approaches. In order to fit the greylevel distribution of the image in parametric approach a statistical model is needed to assume first[3]. A histogram is designed which pictorially represents the set of parameters that control the fitness. By optimizing an objective function the thresholds can be chosen in non-parametric model. The approach that has been proven as robust and more accurate is non-parametric one [4].

A very popular non-parametric thresholding technique is Otsu's method. In this the multilevel thresholding is used to determine various thresholds so that the image can be segmented into several clusters. Otsu's method is an extension to multi-threshold version. If the number of segmented clusters is large then it suffers from a serious drawback [5]. In order to reduce the computational cost for multilevel thresholding several versions of Otsu's method have been developed. The main aim of multilevel thresholding is to segment objects of interest from their background. A non-parametric method that is used for classification and regression in pattern recognition is, the K-Nearest Neighbours algorithm (k-NN) [6]. In this case, k closest training samples are given as input, whereas the output lies on the chosen method i.e., it varied depends on whether k-NN is used for classification or regression. The output can be

assigned to the class as a positive integer. An object is assigned to the single nearest neighbour, if the value of $k=1$. It is a type of instance –based learning or lazy learning and is also a simplest of all machine learning algorithms. In k -NN classification the neighbours are chosen to which the objects property value is known. The main property of k -NN algorithm is Feature extraction. The input data would be transformed into a reduced representation set of features (features vector) if the input data is too large to an algorithm. Feature extraction is defined as an transformation of input data into the set of features. The transformed data in feature space is chosen, in that feature extraction is performed on raw data. One of the best examples for computer vision computation pipeline using k -NN which includes feature extraction and dimension reduction is Haar face detection, Mean-shift tracking analysis, PCA or Fisher LDA projection [7]. In order to estimate the continuous variable in k -NN regression, the k -NN algorithm is used. It uses an weighted average that is derived by the inverse of distance [8]. The steps to implement the algorithm is

- Euclidean distance is computed in prior from the query and is labelled
- Inorder to increase the distance the labelled query is ordered.
- Optimal No. of k nearest neighbours is chosen by using cross validation.
- Finally, an inverse distance is calculated with the k -nearest neighbours.

According to the world health report of the World Health Organization(WHO) in 2004, the most common cause of death is due to Lung cancer. It causes 1.3 million annual deaths worldwide in 2006. The following methods are allowed for early detection of cancer is helical low-dose CT (Computed Tomography) scan, x-ray, and bronchoscopy. Also some other non- invasive imaging techniques such as radio- graphy and computed tomography (CT) is used for detection of lung cancer. Tools such as computer-aided detection/diagnosis (CAD/CADx) are developed in order to support the specialist physician. The lung nodules can be easily identified from digital images so that a series of measurements can be executed [9].

The main aim of Lung nodule image segmentation is to detect lung nodules by the use of computed tomography. Nodules can be visible through X-ray and they are typically asymptotic. The size of the nodule is usually smaller than 3–4 cm in diameter (no larger than 6 cm) and is always surrounded by normal, functioning lung tissue. The intensity ranges from -300 to 0 HU and are fairly common on X-ray images of chest. One of the primary lung cancer tumours from the parts of the body is pulmonary nodules. To obtain the partition efficiently and to segment the structures with features similar to lung nodules an algorithm named the growing neural gas (GNG) is used. Also, a transformation technique known as distance transform which is connected to blood vessels and bronchi is used to divide the segmented features. Finally, to classify the structures in Lung nodules an support vector machine (SVM) is used along with their set of shape and texture features [10].

II. PROPOSED SYSTEM

The main objective of the proposed system is to produce the following:

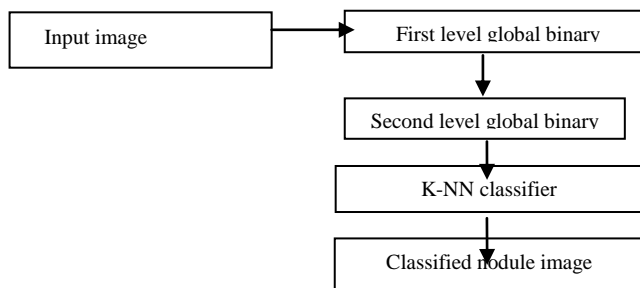


Fig1: Overall System diagram

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization

Volume 4, Special Issue 5, April 2015

International Conference On Emerging Trends in Engineering and Technology (ICETET'15)

On 13th & 14th March 2015

Organized by

Pandian Saraswathi Yadav Engineering College, Arasanoor, Sivagangai, Tamilnadu, India

The algorithm for the global thresholding is as follows

An image $f(x, y)$ and $z(i, j)$ is the grey scale value in pixel. The threshold value is calculated based on the following steps

Step1: Calculate the minimum and maximum of gray scale value with respect to z_i and z_j .

Step2: Initial threshold values is as $T_0 = (z_i + z_j) / 2$.

Step 3: T_k segments the image based on the mean obtained from the grey scale value of Z_0 and Z_B .

$$Z_0 = \frac{\sum_{z(i,j) < T_k} z(i,j) \times N(i,j)}{\sum N(i,j)}$$

$$Z_B = \frac{\sum_{z(i,j) > T_k} z(i,j) \times N(i,j)}{\sum N(i,j)}$$

Step4: Calculate the new threshold $t_{k+1}, t_{k+1} = (z_0 + z_B) / 2$

Step5: if $T_k = t_{k+1}$, end; else $k = k + 1$; goto step 2.

Step6: After the end of step 4, T_k is the best threshold value.

A. FIRST LEVEL GLOBAL THRESHOLDING

CT image is given as input to the first level of global thresholding. The threshold is based on the maximum and minimum gray scale value of an image. The initial threshold value is obtained from the sum of the maximum and minimum gray scale value divide by the 2. The Threshold value is set T value. After the thresholding the input image is segmented as look like in image. This first level thresholding extract the parenchyma regions



B. SECOND LEVEL GLOBAL THRESHOLDING

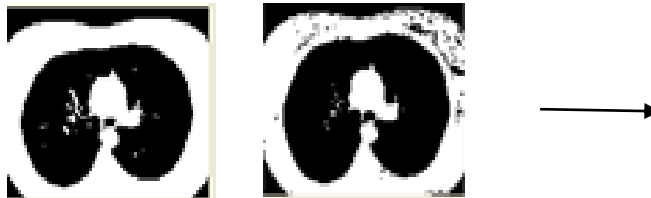
First level o thresholding is given as input to the second level of global; thresholding. The threshold is based on the mean value obtained from the segmented image

$$Z_0 = \frac{\sum_{z(i,j) < T_k} z(i,j) \times N(i,j)}{\sum_{z(i,j) < T_k} N(i,j)}$$

$$Z_B = \frac{\sum_{z(i,j) > T_k} z(i,j) \times N(i,j)}{\sum_{z(i,j) > T_k} N(i,j)}$$

$z(i,j) < T$

The new thresholding value is obtained for the next segmentation from the T_{k+1} value. The T_{k+1} value is calculated until it is equal to the t_k value. The output of the Second level global thresholding It removes all the noises and the trunk portions of the image.



C. K-NN classifier

The output of the second level global thresholding is subjected to classify the presence of lung nodule in the image. The K-nearest neighbour's algorithm is a method for classifying objects based on closest training 1. A positive integer k is chosen K-Nearest Neighbours (KNN) classification divides data into For each row of the test set, the K nearest (in Euclidean distance) a test set and a training set



III. CONCLUSION

The proposed model which uses multilevel thresholding for segmentation. The multilevel thresholding which has two levels of thresholding. The first level threshold image is given a input to the second level thresholding. This multilevel segmenting removes the noise and trunk regions and provides the parenchyma region. This multiple thresholding segmentation of parenchyma region achieves better result than the other thresholding method. The

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization

Volume 4, Special Issue 5, April 2015

International Conference On Emerging Trends in Engineering and Technology (ICETET'15)

On 13th & 14th March 2015

Organized by

Pandian Saraswathi Yadav Engineering College, Arasanoor, Sivagangai, Tamilnadu, India

segmented image is classified based on the K-NN classifier. This classifier classifies the lung nodule based on the Euclidean distance. The distance is a simple calculation method which reduces the cost and runtime. The results show that the proposed method provides better quality for the nodule identification

IV. FUTURE ENHANCEMENT

The proposed scheme only segments and classifies the lung nodule is cancerous or non cancerous. In future the size of the nodule is going to be identified to predict the growth of the nodule. Different classifiers are also employed to classify the nodules much accurate than the K-NN

REFERENCES

- [1] J. J. Erasmus, J. E. Connolly, H. P. McAdams, and V. L. Roggli, "Solitary pulmonary nodules: Part I. morphologic evaluation for differentiation of benign and malignant lesions," *Radiographics*, vol. 20, no. 1, pp. 43-58, 2000.
- [2] B. Zhao, "Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm," *J. Appl. Clin. Med. Phys.*, vol. 4, no. 3, pp. 248-260, 2003.
- [3] Y. Song, W. Cai, and D. Feng, "Microscopic image segmentation with two-level enhancement of feature discriminability," in *Proc. Int. Conf. Digital Image Comput.: Techn. Appl.*, 2012, pp. 1-6.
- [4] M. Brown, D. Aberle, *et al.*, "Model-based segmentation architecture for lung nodule detection in CT," *Radiology* **217**(P), 2000.
- [5] J.K. Leader, B. Zheng, R.M. Rogers, F.C. Sciurba, A. Perez, B.E. Chapman, S. Patel, C.R. Fuhrman, D. Gur, Automated lung segmentation in X-ray computed tomography: development and evaluation of a heuristic threshold-based scheme, *Academic Radiology* 10 (11) (2003) 1224-1236.
- [6] J.R.F. da Silva Sousa, A.C. Silva, A.C. de Paiva, Lung structure classification using 3D geometric measurements and SVM, in: CIARP, Iberoamerican Congress on Pattern Recognition, 2007, pp. 783-792.
- [7] G. Zhu, S. Zhang, Q. Zeng, and C. Wang, "Boundary-based image segmentation using binary level set method," *Optical Engineering*, vol. 46, pp. 050501-050501-3, 2007.
- [8] K. Zhang, L. Zhang, H. Song, and W. Zhou, "Active contours with selective local or global segmentation: A new formulation and level set method," *Image and Vision Computing*, vol. 28, pp. 668-676, 2010. [13]
- [9] N. Rousopoulos, S. Kelley, and F. Vincent. Nearest neighbour queries. In *Proceedings of the ACM SIGMOD Conference*, pages 71-79, 1995.
- [10] B.V. Dasarthy. Nearest Neighbor Norms: NN Pattern Classification Techniques. IEEE Computer Society Press, 1991.