

A Segmentation and Analysis of CT Lung Images Using Graph Cut Technique and Mean Shift Analysis

R.Chinnakannan Ajith

Computer Science&Engineering, VV College of Engineering, Tuticorin, India.

Dr. Ebenezer Juliet

Computer Science&Engineering, VV College of Engineering, Tuticorin, India.

Abstract – Lung segmentation is often performed as a pre-processing step on chest Computed Tomography (CT) images because it is important to identify lung diseases in clinical evaluation. Hence, research on lung segmentation has received much attention. Computer aided diagnosis system provides early lung disease diagnoses. Till today, there are a large number of techniques available that can extract the required foreground from the background. In this project, two efficient segmentation techniques like graph cut and mean shift analysis are studied, implemented and tested with CT lung images. The graph cut based segmentation method has obtained a huge attention because it uses both regional and boundary information. This graph cut segmentation technique uses min-cut algorithm to separate lung from the background. In Mean shift segmentation technique, the bandwidth parameter is estimated using fixed bandwidth estimation. Particle swarm optimization algorithm is used to optimize this bandwidth parameter. Finally, the segmentation results obtained from both the techniques are compared using True Positive Rate and False Positive Rate. The result of segmentation can help the radiologist in early diagnoses of lung diseases. Graph cut segmentation takes a less time for processing but mean shift segmentation takes a more time compare then graph cut. The Both algorithm is also applicable for other parts of human body, such as liver, brain.

Index Terms – Graph cut, Mean shift, GMM, mincut,

1. INTRODUCTION

Image segmentation is the process of partitioning a digital image into different regions which may have similar colour, intensity or texture. Segmentation as a pre-processing step plays a significant role in computer vision, object recognition, tracking and image analysis. Conventionally, segmentation can be grouped into five categories.

- Threshold based segmentation
- Edge based segmentation
- Region-based segmentation
- Watershed based segmentation
- Energy based segmentation

Medical Image Analysis provides a forum for the dissemination of new research results in the field of medical and biological image analysis, with special emphasis on efforts related to the applications of computer vision, virtual reality and robotics to biomedical imaging problems.

An interesting source of images is in the medical field. Here, imaging modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET) etc., generates a huge amount of image information. Not only the size and resolution of the images grow with improved technology, the number of dimension also increases. Previously, medical staff studied two-dimensional images produced by X-ray. Now, three dimensional image volumes are common in everyday practice. Even four dimensional data (three-dimensional images changing over time, i.e. movies) is often used. This increase in size and dimensionality provides major technical challenges as well as cognitive. Automatic, or semi-automatic, algorithms are of interest.

Developing algorithms for medical image analysis requires thorough validation studies to make the results usable in practice. This adds another dimension to the research process which involves communication between two different worlds - the patient-centered medical world, and the computer centered technical world. The symbiosis between these worlds is rare to find and it requires significant efforts from both sides to join on a common goal.

The aim of this paper is to develop a segmentation method for medical imaging applications. In particular, the segmentation of lungs from CT images. The motivation for this work is to increase the segmentation accuracy which is used to increase patient safety by providing better and more precise data for medical decisions, and can directly provide explicit lung regions without any post processing operations even in complex scenarios. First step is to develop a graph cut algorithm, then improve the accuracy by using the Gaussian Mixture Model. This gives only 78% of accuracy over 10 different images and Mean shift analysis method is introduced. This algorithm gives 71% of accuracy over 10 different images.

The segmentation is an initial processing step for clinical diagnosis. Most of the methods in the present are a combination of a variety of methods, and are time-consuming. It cannot get the segmentation result at once, and require multiple steps. This multiple steps are avoided by graph cut algorithm with GMM and Mean shift analysis.

2. METHODOLOGY

The goal of this paper is segmentation of lungs on chest CT images and improving the accuracy. Lung segmentation is the first basic step that segments lung from chest CT images and that will be used for subsequent analysis, such as early lung disease detection or local area analysis. In this work, the accuracy is improved by using graph cut segmentation and Mean shift analysis and then the accuracy obtained from the two techniques is compared. Figure 1 describes this work. Graph cut segmentation and Mean shift analysis are explained below.

3. GRAPH CUT METHOD

The graph cut based segmentation method has obtained a number of attention because this method use both regional and boundary information. This graph cut segmentation technique uses min-cut algorithm to separate lung from the background.

A. GAUSSIAN SMOOTHING

In general, all images will have a noise but only the noiseless images are needed for processing. So the images are involved to a smoothing process. Lot of smoothing techniques is available. It includes Gaussian smoothing, exponential smoothing etc., Gaussian smoothing is suitable for medical images so that they are used in this project. Due to the imaging machinery, environmental conditions, and other external interference, there will be inevitable noise when the CT images are created. This existing noise causes difficulties in analysis, and it affects the segmentation accuracy.

Hence the first step in CT image process is to remove the noise. Commonly used CT image processing methods for noise reduction are mean, median, or Gaussian filtering. To improve the accuracy of segmentation of the lung pulmonary region, the Gaussian filter approach with the Gaussian kernel radius $r=0.5$ is used in this experiment. With that parameter the noise of the CT images can be removed and the contour of lungs can be kept clear.

B. SEED POINTS SELECTION

Seed point is nothing but the pixel values of image. This project needs two seed points, one from background value and the other from object/foreground. Finally all pixels are labeled based on this two pixel values. The value of the pixel is a number that corresponds to the intensity of the image. Seed point selection can be done either manually or automatically but in this project it is done automatically.

i) Generation of GMM

Graph is created by using some technique and then weight is assigned to it. Weight can be assigned by various methods. Every method takes various time, but here by using GMM, processing time is reduced. Each seed point region is modeled using GMM with K sub classes. As lung CT images are gray scale images, the value of the pixel is a number that corresponds to the intensity of the image. To determine the probability model, a mixture of Gaussian distribution of the following form is used.

$$pr(x) = \sum_{k=1}^K \beta_k \cdot N(x; \mu_k, \sigma_k^2) \quad (1)$$

Where K is the number of components, β are weights and $N(\cdot)$ is the Gaussian Probability Density Function (PDF) parameterized by mean value μ_k and standard variance value σ_k^2 :

$$N(\mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left\{-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}\right\} \quad (2)$$

To perform the parameterized learning of priori knowledge mixed model, the EM algorithm is usually the best choice. It not only provides initial robust and strong arguments, but also provides maximum likelihood estimation. Furthermore, compared with the K-means algorithm, the EM algorithm is more suitable for graph cuts, which is just a maximized likelihood segmentation model. Here, the EM-MAP (Maximum a Posteriori) method [1] is used, and the process is defined below.

1. Initialization : $\theta^{(0)}$
 $= \{\beta_1^{(0)}, \dots, \beta_k^{(0)}, \mu_1^{(0)}, \dots, \mu_k^{(0)}, \sigma_1^{(0)}, \dots, \sigma_k^{(0)}\}$

2. E - Step : let $\gamma(i, k) = \beta_k pr(x_i | \mu_k, \sigma_k) / \sum_{j=1}^k \beta_j pr(x_i | \mu_j, \sigma_j)$

3. M - Step : $N_k = \sum_{i=1}^N \gamma(i, j) \dots \dots$

4. Iterate steps 2 and 3 until an arbitrary error is reached:

$e = L(\theta)^{i+1} - L(\theta)^i < \epsilon;$

5. Calculate the final θ^* .

Because EM is a supervised parameter estimation method, the initial values of the model parameter should be given, and the method's result is sensitive to the initial values. Therefore, in the initialization step, the K-means clustering algorithm is used that determines the initial values, including mean value, variance, and weight factors. Using the above EM-MAP

algorithm, the posterior probability for every interesting region in the CT image can be determined. The regional penalty $R_p(\cdot)$ in formula (9) can then be written as follows:

$$R_p(1) = -\ln \Pr(I_p | \text{obj}') \quad (3)$$

$$R_p(0) = -\ln \Pr(I_p | \text{bkg}') \quad (4)$$

Furthermore, the boundary penalty $B\{p, q\}$ in formula (6) uses an ad hoc functions as follows:

$$B_{\{p,q\}} \propto \exp\left(-\frac{(x_p - x_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{dist}(p, q)} \quad (5)$$

The energy function in (8) can then be rewritten as follows:

$$E(A, K, \theta^*) = \lambda \cdot \left\{ -\sum \log(\Pr(A_p | K, \theta^*)) \right\} + \sum B_{\{p,q\}} \cdot \delta(A_p, A_q) \quad (6)$$

With the region penalty and boundary penalty values, the

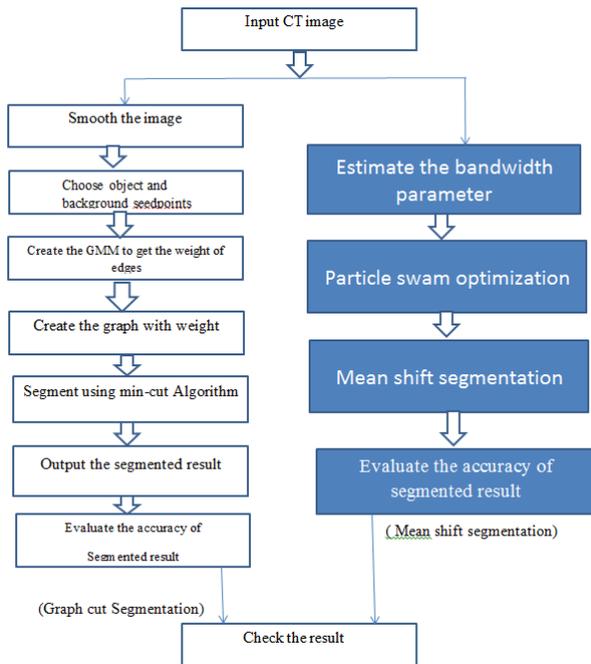


Figure 1 .Overall Process

(As shown in Figure 1) can be created and segmented using graph for the CT image the minimum cut Algorithm.

ii) Graph cut Algorithm

Let an undirected graph be denoted as $G = \langle V, E \rangle$ where V is a series of vertices and E is the graph edge which connect every two neighbor vertices. The vertex V is composed of two different kinds of nodes (vertices). The first kind of vertices is neighborhood nodes which correspond to the pixels and the other kind of vertices are called terminal nodes which consist of s (source) and t (sink). This kind of graph is also called s-t

graph where, in the image, s node usually represents the object while t node denote the background.

In this kind of graph, there are also two types of edges. The first type of edges is called n links which connect the neighboring pixels within the image (Here adopt 4-connected system in the 2D image). And the second type of edge is called t -links which connect the terminal nodes with the neighborhood nodes. In this kind of graph, each edge is assigned with a non-negative weight denoted as w_e which is also named as cost. A cut is a subset of edges E which can be denoted as C and expressed as $C \subseteq E$. The cost of the cut $|C|$ is the sum of the weights on edges C which is expressed as follows [4].

$$|c| = \sum_{e \in C} W_e \quad (7)$$

A minimum cut is the cut that have the minimum cost called min-cut and it can be achieved by finding the maximum flow. The graph is divided by this cut and the nodes are separated into two disjoint subsets S and T where $s \in S, t \in T$ and $S \cup T = V$. The two subsets correspond to the foreground and background in the image segmentation. This kind of graph is depicted in Figure 2. The image pixels correspond to the neighbor nodes in the graph (except s and t nodes). The solid lines in the graph are n -links and the dotted lines are t -links.

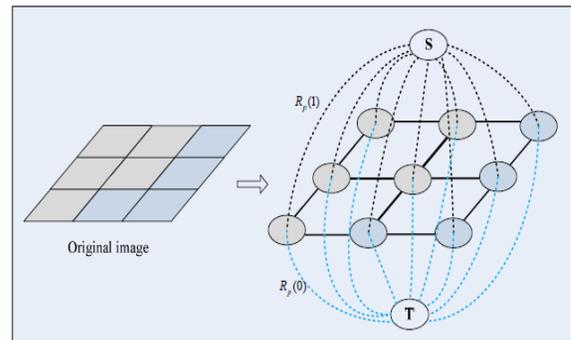


Figure 2. Illustration of s-t graph.

The image pixels correspond to the neighbor nodes in the graph (except s and t nodes). The solid lines in the graph are n -links and the dotted lines are t -links.

iii) Graph cut Segmentation

Image segmentation can be regarded as pixel labeling problems. The label of the object (s -node) is set to 1 while that of the background (t -node) is set to 0 and this process can be achieved by minimizing the energy-function through minimum graph cut. In order to make the segmentation reasonable, the cut should be occurred at the boundary between object and the background. At the object boundary, the energy (cut) should be minimized. Let $L = \{l_1, l_2, l_3, \dots, l_i, \dots, l_p\}$ where p is the number of the pixels in the image and $l_i \in \{0, 1\}$. Thus, the set L is divided into two parts and the pixels labeled with 1 belong to object

while others are grouped into background. The energy function is defined by the following equation.

$$E(L) = \alpha R(L) + B(L) \tag{8}$$

Where, $R(L)$ is called regional term which incorporates the regional information into the segmentation and $B(L)$ is called boundary term which incorporates the boundary constraint into segmentation, α is the relative importance factor between regional and boundary term. When α is set to 0, it means that the regional information is ignored and only the boundary information is considered. In the energy function in eq. (8), the regional term is defined by the following equation.

$$R(L) = \sum_{p \in P} R_p(l_p) \tag{9}$$

Where, $R_p(l_p)$ is the penalty for assigning the label l_p to pixel p . The weight of $R_p(l_p)$ can be obtained by comparing the intensity of pixel p with the given histogram (intensity model) of the object and background. The weight of the t-links is defined by the following equations

$$R_p(1) = -\ln \Pr(l_p | \text{obj}') \tag{10}$$

$$R_p(0) = -\ln \Pr(l_p | \text{bkg}') \tag{11}$$

From eq. (4) and (5), it can be noted that when $\Pr(l_p | \text{obj}')$ is larger than $\Pr(l_p | \text{bkg}')$, $R_p(1)$ will be smaller than $R_p(0)$. This means when the pixel is more likely to be the object, the penalty for grouping that pixel into object should be smaller which can reduce the energy in eq. (8). Thus, when all the pixels have been correctly separated into two subsets, the regional term would be minimized. $B(L)$ in eq. (8) is the boundary term which is defined by the following equation

$$B(L) = \sum_{\{p,q\} \in N} B_{\langle p,q \rangle} \cdot \delta(l_p l_q) \tag{12}$$

Where p, q is neighboring pixels and $\delta(l_p l_q)$ is defined as:

$$\delta(l_p l_q) = \begin{cases} 1 & \text{if } l_p = l_q \\ 0 & \text{if } l_p \neq l_q \end{cases} \tag{13}$$

For the regional constraint, it can be interpreted as assigning labels l_p, l_q to neighboring pixels. When the neighboring pixels have the same labels, the penalty is 0 which means the regional term would only sum the penalty at the segmented boundary. For the term $B_{\langle p,q \rangle}$, it is defined to be a non-increasing function of

$|I_p - I_q|$ as follows [4]

$$B_{\langle p,q \rangle} \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \tag{14}$$

Where σ can be viewed as camera noise. When the intensity of two neighboring pixel is similar, the penalty is very high. Otherwise, it is low. Thus, when the energy function obtains minimum value, it is more likely to be occurred at the object boundary.

The weight of the s-t graph is given as following.

$$\text{Weight} = \begin{cases} B_{\langle p,q \rangle} & \{p, q \in \text{Neighboring pixel}\} \\ \alpha \cdot R_p(0) & \text{for Edge } \{p, S\} \\ \alpha \cdot R_p(1) & \text{for Edge } \{p, T\} \end{cases} \tag{15}$$

Eq. (15) can also be explained as that, in the s-t graph, when the intensity of the pixel is inclined to be the object, the weight between this pixel and s-node will be larger than that between pixel and t-node which means the cut is more likely occurred at the edge with smaller weight.

For the neighboring pixels, when their intensity is very similar, the weight is very big which is not likely to be separated by the cut. Thus, when the minimum cut is achieved from the s-t graph, the location of the cut is close to the object boundary. The implementation of the graph cut can be fulfilled by the max-flow/min-cut. Figure 3. illustrates the graph cut for 3x3 image segmentation. The cut corresponds to the minimal energy of eq. (2)

B) Mean shift method

Mean shift is a nonparametric estimator of density. In medical analysis of pulmonary CT images, the detected lung areas segmentation has differences in image colour. The advantage of this method from other segmentation method is that, the number of segments need not be determined but in exchange we have to estimate a fit amount for bandwidth. In this project work PSO algorithm is used for estimating the bandwidth parameter. This method is more efficient and accurate from other methods.

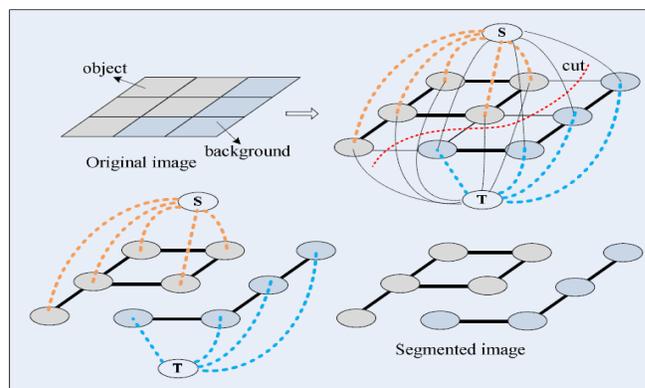


Figure 3. Illustration of graph cut for image segmentation.

4. MEAN SHIFT METHOD

Mean shift is a nonparametric estimator of density. In medical analysis of pulmonary CT images, the detected lung areas segmentation has differences in image colour. The advantage of this method from other segmentation method is that, the number of segments need not be determined but in exchange

we have to estimate a fit amount for bandwidth. In this project work PSO algorithm is used for estimating the bandwidth parameter. This method is more efficient and accurate from other methods.

A) Mean shift analysis

Mean shift algorithm has the ability to cluster points in one space around special points that determines the center points. This process of determining centers and clustering point based on gradient density function is continuously done before the points have stationary place. The function that places points around the centers is based on the similarity computation between centers and points gathered in an outline, named membership function. This process of choosing member from the similarities in color or intensity can be used for removing noise or other mismatching functions. The image space is segmented by empirical probability density functions (PDF). Mean shift is a procedure that gives us discrete sample of density function. It is useful for detecting the modes of this density. This is an iterative method that starts with an estimative initialization and then uses kernel density estimation function. Kernel density estimation (KDE) is a non-parametric way to estimate the probability density function of a random variable. The kernel density estimation is used to resolve smoothing problems with different smoothing bandwidths [11]. For the definition of kernel density estimation, let (x_1, x_2, \dots, x_n) be the discrete samples drawn from some distribution with an unknown density f . Schema off is estimated and then the density is computed with,

$$f_H(x) = 1/n \sum_{i=1}^n K_H(x - x_i) \tag{16}$$

Where $K(\cdot)$ is the kernel function, n is the number of samples and the smoothing parameter $h>0$ is a bandwidth parameter and it is defined by,

$$K_H(x) = |H|^{-\frac{1}{2}} K(H^{-\frac{1}{2}} x) \tag{17}$$

H is the positive definite bandwidth matrix. For correlation between the proposed method for estimating kernel density and the real kernel density, mean square error (MSE) criteria are used. The asymptotic approximation is one of the best choice. The computation is minimized by using Epanechnikov kernel:

$$K_E(x) = \begin{cases} 1-x & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases} \tag{18}$$

This shows radially symmetric kernel by $KE(x)$.

$$K_E(x) = \begin{cases} \frac{1}{2} C_d^{-1} (d+2)(1-\|x\|^2) & \|x\| \leq 1 \\ 0 & otherwise \end{cases} \tag{19}$$

C is the volume of the unit d -dimensional sphere and parameter d is the number of dimensions.

$$K_N(x) = e^{-\frac{1}{2x}}, x \geq 0 \tag{20}$$

The normal kernel is denoted by $KN(x)$.

$$K_N(x) = (2\pi)^{-\frac{d}{2}} (c^{-\frac{1}{2}} \|x\|^2) \tag{21}$$

Equations (19) and (21) are suitable for most of the applications. It is assumed that $H=h^2$ then,

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) = \frac{1}{nh} \sum_{i=1}^n K_h\left(\frac{x-x_i}{h}\right) \tag{22}$$

The kernel is divided into two different radially symmetric kernels (d , spatial domain and r , color range). Then the kernel density function is:

$$\hat{f}_h(x) = \frac{c}{n(h^d)^p (h^r)^q} \sum_{i=1}^n K_h^S\left(\left\|\frac{x^d-x_i^d}{h_i^d}\right\|^2\right) K_h^r\left(\left\|\frac{x^r-x_i^r}{h_i^r}\right\|^2\right) \tag{23}$$

In (23) for an image data, p shows spatial domain dimension ($p=2$) and d shows color domain ($d=3$).

B) Particle Swarm optimization

This algorithm was first represented in 1995 by Kennedy and Eberhat as a heuristic algorithm to solve optimization problems and was first intended for simulating social behaviors a stylized representation of the movement of organisms in a bird flock or fish school. Different algorithms of PSO have been classified by their optimization in continuous, fuzziness, matching, convergence speed, topology, aggregation, mobility, particle distribution function, quantum particles, and etc. [13]. This algorithm used the topology of global best PSO and the topology of local best PSO simultaneously. The global best topology derived each particle into the best point that found in the group while the local best topology targeting the best experience of the same particle. The resultant of these orientations is the vector that is considered as the area under the curve ROC and that should be maximized [12]. First we consider the X matrix as an initial random population in which each row corresponds to one particle. The number of objective function parameters, determine each particle dimension that is equal to the number of column of X matrix. In the proposed algorithm, PSO is used for estimating bandwidth parameter in mean shift kernel estimation. Each particle is considered as a random result for the problem. In each execution of the algorithm, the best experience of each particle, the best republican experience stored and the new rate of each particle is obtained from eq. (24).

$$V_i(t) = wV_i(t - 1) + c_1 Rand_1(p_{i,best} - X_i(t)) +$$

$$c_{21}Rand_{21}(p_{g,best} - X_i(t)) \tag{24}$$

$$X_i(t) = X_i(t - 1) + V_i(t) \tag{25}$$

In (24) and (25), V is the speed of particle, w is the weight coefficient of speed, $Rand$ is the random number in range (0-1), c_1 is the coefficient of local rate, c_2 is the coefficient of global rate, $P_{i,best}$ is the best experience of each particle, $P_{g,best}$ is also the best republican experience, X matrix is the position of all particles. The value of w is decreased in each repeat of algorithm with parameter attenuation coefficient until the bound limitation become smaller within algorithm iterations.

C) Bandwidth kernel density estimation

Mean shift is one of the most useful methods of segmentation and tracking. However, one of the problems in this method that can cause the limitation of this procedure is the determination of its scale parameters. In this section, the most popular bandwidth density estimators and their advantage are discussed. These methods are divided into two categories, fixed bandwidth and variable bandwidth. Variable bandwidth consists of two methods, the balloon and the sample point.

The fixed bandwidth kernel estimate is defined in eq. (26),

$$f_n(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right), x \in R^d \tag{26}$$

Where sample x_i is chosen from some unknown density function f and kernel K is defined in R^d . In fixed bandwidth method, h is defined as a constant across $\forall x \in R^d$. The density of each point can be estimated from eq. (11). The best value for h parameter is the value that correlates between $f(x)$ and $f(x)$. It is determined by the mean squared error (MSE) criteria. The MSE is defined as the sum of the variance and squared bias as shown in (27).

$$MSE(x) = E[\hat{f}(x) - f(x)]^2 = Var(\hat{f}(x)) + (Bias(\hat{f}(x)))^2$$

$$Var(x) \approx n^{-1}h^{-d}R(K)f(x)$$

$$Bias(x) \approx \frac{1}{2}h^2\mu_2(K)\Delta f(x) \tag{27}$$

For minimizing MSE a tradeoff between bias and variance is needed as shown in (27). The bias is proportional to h^2 , and then increasing h will result in an increase in bias. Similarly, decreasing h will lead to decrease in variance. Therefore an optimum value of h is achievable by minimizing the mean integrated squared error (MISE) as shown in (28).

$$MISE(x) = E \int (\hat{f}(x) - f(x))^2 dx \tag{28}$$

Instead of using a fixed value for bandwidth parameter, it is better to choose an adaptive parameter by substituting $h(x)$ with h . Then $h(x)$ can estimate bandwidth for each point x . In other

words, two different values for h are computed: one for pixels that are located in the neighborhood of zero and one for pixels that located in higher range. In balloon estimation method, f is estimated at each point x by calculating the average of scaled kernels centered at each data point. The balloon estimator is calculated using equation (29).

$$\hat{f}_1(x) = \frac{1}{nh(x)^d} \sum_{i=1}^n k\left(\frac{x-x_i}{h(x)}\right) \tag{29}$$

Balloon estimator has significant improvements when compared with the fixed bandwidth method. In this method, bias and variance are proportional to h^2 and usually fails to integrate to one.

D) Using PSO using for determining bandwidth parameter

In this section, by tuning the Particle Swarm Optimization (PSO) algorithm parameters, the bandwidth of mean shift method is estimated. Fixed bandwidth estimation method is used for estimating parameter h in mean shift algorithm. Achieve optimum value for parameter h , MISE should be minimized in eq. (12). At first, an estimation of $f(x)$ with a constant h is used to determine $nh(x)$. In PSO algorithm each particle has one parameter that optimizes the fitness function of mean shift. An AMISE criterion that has acceptable accuracy in parameter estimation is used. Figure 4 shows the pseudo-code of proposed algorithm.

The parameter of PSO algorithm is shown in Table I.

Table I. Parameter of PSO Algorithm

<p>Input: Fitness function, Number of particles, Number of parameters per particle, velocity coefficient, weight coefficient, damp ratio, parameter range</p> <p>Output: Best input the optimized AMISE criteria</p> <ol style="list-style-type: none"> 1. Fetch data, initialize population of PSO 2. For j=1 to maximum iteration <ol style="list-style-type: none"> a. New random value, calculate weight coefficient 'W' b. Calculate new velocity vector c. Calculate new coordination d. If new coordination is over range, set it to boundary value e. Fitness evaluation f. If new criteria is better than previous value <ol style="list-style-type: none"> Then (i)Replace better value as global best (ii)Replace better value as local best g. End if 3. End Loop
--

Figure 4. The Pseudo code of evaluated PSO algorithm

5. EXPERIMENTAL RESULTS

a) EXPERIMENTAL SETUP

The data of chest CT images that is used to evaluate the performance of the proposed algorithm are collected from various websites. The format of the data images are of various sizes like 190*200, 271*186 and etc. This project is implemented in MATLAB R2013a and the supported system is Intel® Pentium® CPU A1018@2.10GHZ,2GBRAM. The experiments investigate the effect of performance on the following aspects.

b) QUANTITATIVE EVALUATION

To evaluate the performance of the segmentation approach, only True Positive Rate (TPR) & False Positive Rate (FPR) are used. TPR is the proportion of positive cases that were correctly identified (16). The FPR is the proportion of Negative cases that were incorrectly classified as positive (17). Then find out the accuracy of segmented image. Accuracy is calculated using equation (18).

$$TPR = TP / (TP + FN) \tag{16}$$

$$FPR = FP / (FP + TN) \tag{17}$$

Parameter	Variable	value
Weight of coefficient	w	0.5
Local velocity coefficient	C1	2
Global velocity coefficient	C2	2
No. of Particles	NP	10
Max.Iteration	Iter	100
Damp Ratio	DR	0.95
Lower Bound	LB	0
Upper bound	UB	100

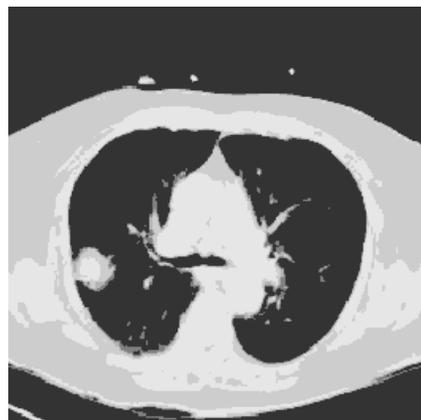
$$Accuracy = ((TP + TN) / (TP + FP + FN + TN)) \tag{18}$$

Where TP is True Positive, FP is False Positive, FN is False Negative, TN is True Negative.

Segmented Result is shown following



Graph cut algorithm



Mean shift algorithm

6. CONCLUSION

The importance of lung segmentation in lung CT image processing and the clinical analysis of lung disease, research on lung segmentation has received much attention in the past and many segmentation algorithms have been proposed. To facilitate the computer-aided lung lesion treatment planning, quantitative assessment of lung cancer treatment response and robust lung segmentation method is also needed. The graph cut algorithm has achieved spectacular progress in computer vision and image processing field, which inspires us to investigate its power in lung segmentation of CT images. In this work, a new method based on graph cut algorithm improved with GMMs is developed. The experimental results showed that the proposed method can directly provide explicit lung segmentation without any post-processing such as morphological operation. All the cost mainly focuses on GMMs training process, the construction of corresponding graph and solving of the minimum-cut algorithm. This Algorithm is compared with a mean shift algorithm. Mean shift method uses adaptive mean shift method that estimate the bandwidth parameter by using fixed bandwidth estimation. Use Particle Swarm Optimization algorithm to optimize band width parameter. The large set of lung images have shown that the proposed method give better segmentation results when compared to graph cut method and also it present an accurate method to obtain best bandwidth parameter. This method is a good alternative for old estimation and trial and error methods. The both developed algorithm can help us to automate lung CT images as the first step in analysing and diagnosing pulmonary diseases.

REFERENCES

[1] Shuangfeng Dai , Kailua, JiyangDong , YifeiZhang , YongChen : A novel approach of lung segmentation on chest CT images using graph cuts, Elsevier, Neurocomputing, 2015, pp.799-807.
 [2] Lin-Yu Tseng, Li-Chin Huang: An Adaptive Thresholding Method for Automatic Lung Segmentation in CT Images,, IEEE AFRICON, 2009, pp.23-25.

- [3] Yuri Y. Boykov, Marie-Pierre Jolly : Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images, IEEE Transation on Medical Imaging, 2001, pp.105-112.
- [4] Faliu Yi, Inkyu Moon: Image Segmentation: A Survey of Graph-cut Methods, International Conference on Systems and Informatics, 2012, pp.1936-1941.
- [5] Nihad Mesanovic , Mislav Grgic , Haris Huseinagic , Matija Males , Emir Skejic, Muamer Smajlovic : Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm, IEEE Transation on Medical Imaging, 2006, pp.395-400.
- [6] Sema CANDEMIR, Yousf Sinan AKGUL : Statistical significance based graph cut regularization for medical image segmentation, TUBITAK in Elec eng & Comp Sci, vol 19 , 2011, pp.961-972
- [7] Tahir Sa, Mehmet C: Color image segmentation based on multiobjective artificial Bee colony optimization, Elsevier, Applied Soft Computing, 2015, pp.389-401.
- [8] Yuri Boykin, Vladimir Kolmogorov : An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision, IEEE on Pattern Analysis and Machine Intelligence, 2004, pp.1124-1137.
- [9] Shanhui Sun, Christian Bauer, and Reinhard Beichel : Automated 3-D Segmentation of Lungs With Lung Cancer in CT Data Using a Novel Robust Active Shape Model Approach , , IEEE Transation on Medical Imaging, 2012, pp.449-460.
- [10] Amal A. Farag, Hossam E. Abd El Munim, James H. Graham, Aly A. Farag : A Novel Approach for Lung Nodules Segmentation in Chest CT Using Level Sets, IEEE Transation on Medical Imaging, 2013, pp.5202-5213.
- [11] Saeid Fazli, Mitra Jafari, Amir Safaei : Automated Lung CT Image Segmentation Using Kernel Mean Shift Analysis, 8th Iranian Conference on Machine Vision and Image Processing, 2013, pp.392-396.
- [12] Huiyu Zhou, Gerald Schaefer, Abdul H. Sadka, M. Emre Celebi, : Anisotropic Mean Shift Based Fuzzy C-Means Segmentation of Dermoscopy Images, IEEE Journal of Selected Topics in signal processing, 2009, pp.26-34.
- [13] Stefano Diciotti, Simone Lombardo, Massimo Falchini, Giulia Picozz, Mario Mascalchi : Automated Segmentation Refinement of Small Lung Nodules in CT Scans by Local Shape Analysis, IEEE Transaction on Biomedical Engineering , 2011, pp.3418-3428.