

# Segmentation of Parotid Lesions in CT Images using Wavelet-based Features

Tung-Ying Wu

Dept. of Electrical Engineering  
National Chiao Tung University  
Center of Measurement and Standard  
Industrial Technology Research Institute  
Hsinchu, Taiwan (R.O.C)

Sheng-Fuu Lin

Dept. of Electrical Engineering  
National Chiao Tung University  
Hsinchu, Taiwan (R.O.C)

## ABSTRACT

Automatic segmentation of parotid glands for computer-aided diagnosis in clinical practice is still a challenging task, especially when there are lesions needing to be outlined. In the applications of image-based diagnosis and computer-aided lesion detection, image segmentation is an important procedure. Features extracted from image analysis in companion with image segmentation algorithms are used to provide region-based information for clinical evaluation procedures. In this paper, we describe a method for segmenting the parotid regions with skeptical lesions in the head and neck CT images. At first, a modified discrete wavelet transform algorithm, is introduced to decompose an image into sub-bands, and the feature descriptors effective for soft tissues characteristics are computed using the derived coefficients in the sub-bands. Then, clustering algorithms are proposed to connect the pixels corresponding to similar features into several regions of the soft tissues, and so do the tissues of the lesions. In this paper, a comparative study of feature-based segmentation with three methods is carried on, and the extracted regions are compared with the segmentation from the experts for evaluating the performance.

## General Terms

Image processing, Medical image analysis, Image segmentation.

## Keywords

Parotid, Wavelet, Computer Tomography, Image segmentation.

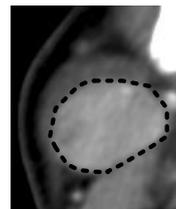
## 1. INTRODUCTION

As image-based computer-aided diagnosis (CAD) is introduced to clinical applications, automatic detection of the anatomical structures, including bones, organs, gland tissues and even tumors in the medical images, is also in high demands and developed in recent years [1-5]. In order to find the targets accurately in images filled with complex anatomical structures, image segmentation plays an important role in these works. Image segmentation methods are proposed to divide an image into several regions according to the image features, and each extracted region can be further processed for diagnosis assistance or guidance using morphological and geometric methods. In head and neck (H&N) CT images, automatic or semi-automatic segmentation methods aiming at soft tissue structures in the CT slices are in

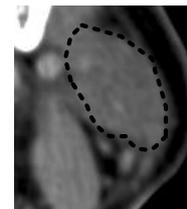
attentions in recent years, because manual segmentation requires skilled techniques and becomes tedious and time-consuming as the amount of data grows. It is beneficial for assisting clinical diagnosis and advanced therapy planning like intensity-modulated radiotherapy (IMRT)[6-7].



(a)



(b)



(c)

**Figure 1: (a) An H&N CT slice with parotid glands at both lateral sides within the dashed rectangles. (b) The left parotid gland in (a) with the lesion tissues inside the boundary of the dashed curve. (c) The right parotid in (a) without lesions.**

However, there are still difficulties in automatically extracting the parotid glands, and the lesions on risk adjacent to the tissues in the CT images. As shown in Figure 1, the parotids locate at both lateral sides of the neck anatomical structures. A skeptical lesion is found at the left parotid and contoured in Figure1.(b). In recent years, atlas-based methods are proposed based on the patterns manually sketched out by the experts [7-11]. However, the stored atlas patterns need previous evaluation to choose an appropriate one because of the variation of the anatomical structures between patients, but

the pattern evaluation in this stage still needs manual procedures on extracting the rough regions of the targets. In order to capture the region-based information for improving medical automation, image segmentation can be an effective method. Besides, the lesions or the swollen glands are not predictable in size or location appeared in the CT images, so detection assisted with image segmentation can make use of more efficient information and prevent unsatisfied results. However, the gray level distribution, which is frequently used in image segmentation, is not sufficient enough for the regions of the parotid tissues with low contrast. Generally speaking, organs are expected to have consistent features shown in the CT slices within the tissues, so that feature-based methods in companion with segmentation algorithms are considered, and have been applied to segment the soft tissues like the livers in abdominal CT images and the brain tissues [14-17]. Therefore, soft tissues in the parotid regions need appropriate feature descriptors for segmentation. To derive the local features in images, wavelet transform is a method frequently mentioned in many applications for deriving the features related to texture analysis [18-20]. The wavelet-based features are characterized by the local statistical properties implying the relationship between the neighbor pixels in accordance with the wavelet coefficients, and as a result, the segmentation method utilizing the derived features can cluster similar pixels into regions based on the local image characteristics.

In this paper, we proposed a method with local region-based features to segment the soft tissues of the parotids and the regions in risk in the region of interest (ROI). At first, the image-based features are derived based on the translation-invariant wavelet analysis, and then the image feature descriptors are derived from the wavelet coefficients and utilized for image segmentation. Several segmentation algorithms utilizing the derived image features are compared in the experiments, and the results are evaluated by comparing with the regions figured out by the clinical experts.

## 2. Wavelet-based Texture Analysis

Wavelet-based texture analysis includes two steps: wavelet transform computation and feature extraction. A modified wavelet transform with translation preservation is mentioned in this section, and the wavelet coefficients are then used to derive the image features corresponding to each pixel.

### 2.1 Wavelet Transform

The wavelet analysis performs multi-scale properties in capturing both location and frequency, and is also proven appropriate for deriving the image features such as the local texture properties [21-23]. The basic idea of the continuous wavelet transform (CWT) of a one-dimensional signal  $f(x)$  can be expressed by the equation shown as below:

$$W_{\psi}(s, \tau) = \int_{-\infty}^{+\infty} f(x) \psi_{s, \tau}(x) dx \quad (1)$$

where  $\psi_{s, \tau}$  means a wavelet :

$$\psi_{s, \tau}(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x - \tau}{s}\right) \quad (2)$$

and  $s$  and  $\tau$  are the parameters related to the scale and translation. In applications,  $s$  varies to realize the multi-scale analysis. As to the analysis on digital signals, the wavelet transform based on discrete wavelet packages are proposed to

decompose a digital signal into sub-bands. A wavelet package is consisted of two bi-orthogonal filters, including a high-pass filter used for deriving the wavelet coefficients and a low-pass filter for the approximate coefficients. For digital signals, the digital wavelet transform (DWT) can be implemented by the inner product equations modified from CWT:

$$c_j(k) = \left\langle f(n), 2^{-j} \phi(2^{-j} n - k) \right\rangle \quad (3)$$

$$w_j(k) = \left\langle f(n), 2^{-j} \psi(2^{-j} n - k) \right\rangle \quad (4)$$

where  $j$  is the dyadic multi-resolution analysis parameter,  $\phi$  is the scaling function,  $\psi$  is the wavelet function, and  $c_j(k)$  and  $w_j(k)$  are the pair of the approximation coefficients and the wavelet coefficients. A digital signal  $f(n)$  can be decomposed into two sub-bands by measuring the similarity with the functions  $\phi$  and  $\psi$  using the inner product computation in equation (3)-(4). The one-dimensional wavelet transform can be extended to the two-dimensional transform by committing the convolution of the wavelet filters with the input signal in both the directions along the column and the row.

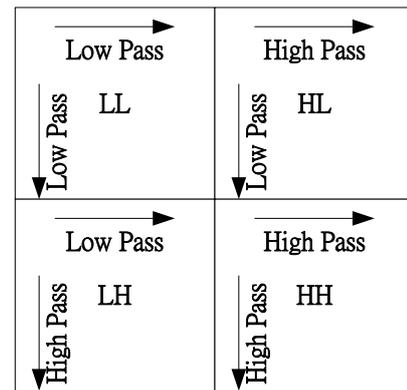


Figure 2. : Four sub-bands after traditional DWT of an image.

As shown in Figure 2, after DWT of an image at a single level, the input image is consequently decomposed into four sub-bands: LL, LH, HL, and HH. The LL sub-band contains the scaling information, and the other three sub-bands, LH, HL, and HH own the detail information of the three directions, including the horizontal, vertical and diagonal. Each sub-band can be in cascades decomposed into another sub-bands in higher scales. The multi-scale analysis in DWT is realized with decimation by 2 after the convolution computation expressed as  $2^j$  in equation (3) and equation (4), but the decimation can not preserve translation invariance and half resolution that may also cause the problem of accuracy. Therefore, a decomposition scheme called as "à trous" (with hole) algorithm is proposed for the drawback [24], and the decimation process in the traditional DWT is replaced with up-sampling the filters by inserting zeros between each coefficients to enlarge the size of the filter. As in Figure 3, the original filter is inserted with zeros between the adjacent entries, and the scale of the filter can be extended and the filter size is also doubled.

$$\begin{matrix} f[-5] & f[-4] & f[-3] & f[-2] & f[-1] & f[0] & f[1] & f[2] & f[3] & f[4] & f[5] \\ \hline f[-5] & 0 & f[-4] & 0 & f[-3] & 0 & f[-2] & 0 & f[-1] & 0 & f[0] & 0 & f[1] & 0 & f[2] & 0 & f[3] & 0 & f[4] & 0 & f[5] \end{matrix}$$

**Figure 3: An example of a filter and the filter inserted with zeros.**

The decomposition via the “à trous” algorithm can generate sub-bands of the same size as the input signal to keep the translation invariance in each scale. However, the wavelet coefficients in the “à trous” algorithm are not directly derived by convolving the signal with the wavelet function as shown in equation (4), but resulting from subtracting the approximation coefficients in two successive scales as equation (5):

$$w_j(k) = c_{j-1}(k) - c_j(k). \quad (5)$$

where  $c_j(k)$  is the approximation coefficients derived from convolving the input signal with the scaling function  $\phi$  at level  $j$ .  $c_j(k)$  and  $c_{j-1}(k)$  are approximation coefficients in successive scales at level  $j$  and level  $j-1$ .

$$c_j(k) = c_{j-1}(k) \otimes \phi_{j-1}. \quad (6)$$

As to a 2D image, the approximation sub-band at level  $j$  can be generated by convolving the input image with the scaling function along the column and row directions:

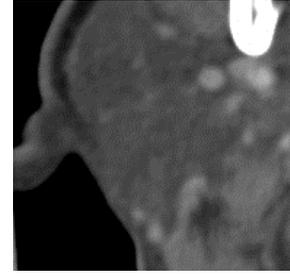
$$c_j(k) = c_{j-1}(k) \otimes_{\rightarrow} \phi_{j-1} \otimes \phi_{j-1} \downarrow. \quad (7)$$

where  $\otimes_{\rightarrow}$  and  $\otimes \downarrow$  denote the convolution operations along the row and column directions at level  $j-1$ , respectively. The decomposition can be illustrated as Figure 4.

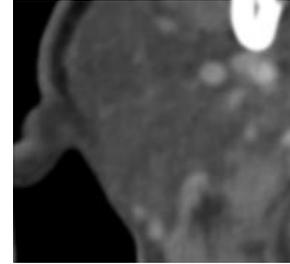


**Figure 4: The decomposition scheme of the “à trous” algorithm.**

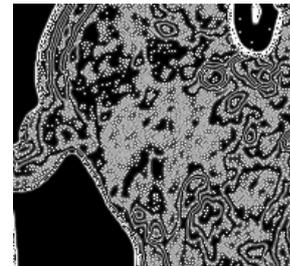
Figure 5 shows an example of a one-level decomposition using the “à trous” algorithm of a parotid CT image. Figure 5.(a) is the original image and Figure 5.(b) illustrates the approximation coefficients deriving from filtering Figure 5.(a) with a low-pass filter in both directions. Besides, Figure 5.(c) illustrates the wavelet coefficients corresponding to each pixel with brightness and the higher brightness indicates the higher intensity of the variation. It can be seen that the wavelet coefficients can reflect the intensity of the regional detail information such as edges.



(a)



(b)



(c)

**Figure 5: (a) The ROI of the parotid in a CT slice (b) image filtered by a low-pass filter (c) the intensity map of the difference between (a) and (b).**

## 2.2 Wavelet-based Feature Extraction

After the DWT computation, the coefficients in the decomposed sub-bands are proposed to derive the feature descriptors. The segmentation of the soft tissues highly depends on the features selected to express the characteristics. There are several descriptors in accordance with these coefficients from DWT, including the energy, entropy, mean, variance, and the contrast, etc. As equation (8) to (10) written below, the measures collect the coefficients inside a mask to compute the relationship between the center pixel of a mask and its neighbor pixels.

$$Energy = \frac{1}{N^2} \sum_{i,j \in mask} |W_{i,j}|^2 \quad (8)$$

$$Entropy = -\frac{1}{N^2} \sum_{i,j \in mask} |W(i,j)|^2 \times \log(|W(i,j)|^2) \quad (9)$$

$$Mean = \frac{1}{N} \sum_{i,j \in mask} W(i,j) \quad (10)$$

$$Variance = \frac{1}{N^2} \sum_{i,j \in mask} |W(i,j) - Mean|^2 \quad (11)$$

where  $W_{i,j}$  means a coefficient deriving from DWT at  $(i, j)$  inside a mask within a sub-band, and  $N$  is the number of elements inside the mask. Energy is a frequently used feature used to measure the variation inside a mask in the sub-bands. Energy in sub-bands implies the variation in directions or scales, and becomes larger at the regions having more detail information. Therefore, it can be used to evaluate the edge intensity within the mask. Besides, entropy is another wavelet texture feature descriptor used to demonstrate the randomness or uniformity in a region. The mean and variance are the statistical features. In conclusion, energy, entropy, and variance derived from the sub-bands of wavelet coefficients can be utilized to evaluate the distinctive variations of each element, and are beneficial for the image segmentation methods using the features related to only the gray level distribution.

The feature vector used in this work is composed of some of the descriptors mentioned above. Besides the approximation coefficients, energy, entropy and variance from the wavelet coefficients are also introduced, and each element of the vector is adjusted by a weighting coefficient.

### 3. Feature-based Image Segmentation

Clustering algorithms used in image process and computer vision are intended to aggregate the pixels with similar features into clusters, and the image can be separate into several regions for advanced analysis. Three segmentation methods are applied in this work after the texture features are extracted, including mean-shift, fuzzy C-means and K-means.

#### 3.1 Mean-Shift Algorithm

Mean-shift algorithm is an intuitive method used to cluster points with the restriction of a determined bandwidth [25][26]. The algorithm needs not define the number of the clusters before segmentation. A kernel density estimation function involving the bandwidth parameter is given to evaluate the weight of a point belonging to a cluster. A Gaussian function expressed as below is usually adopted as an estimation kernel.

$$K(\|x_i - x_c\|) = e^{-c\|x_i - x_c\|^2}, \quad (12)$$

where  $x_i$  denotes a data point,  $x_c$  is the estimated center of a cluster and  $c$  is a positive constant used to adjust the bandwidth.

Mean-shift clustering iteration proceeds with moving the centers of the clusters, and the distance from the cluster center to each data point is computed repeatedly and weighted by the estimation kernel to update the cluster centers. A cluster center is re-estimated by

$$x_c^{t+1} = \frac{\sum K(\|x_i - x_c^t\|)x_i}{\sum K(\|x_i - x_c^t\|)}, \quad (13)$$

where  $x_c^t$  means the center at the  $t$ th iteration.

The procedure terminates while the centers of the clusters do not vary or vary in only a small range and all points are as a result assigned to clusters. Each cluster resulted from the mean-shift algorithm owns the points within the bandwidth.

However, the bandwidth for segmentation is hard to decide in applications and is always obtained from experiment.

#### 3.2 Fuzzy C-Means Algorithms

Fuzzy C-Means (FCM) is an algorithm of soft clustering [27][28]. The idea of the fuzzy logic is employed in this method, and each point can be allowed to belong to two or more clusters. The degree of a point belonging to a cluster is evaluated from 0 to 1 with a smooth function. Different from the mean-shift algorithm, the number of clusters should be previously defined before the fuzzy partitioning. Mathematically, FCM is based on minimization of the objective function:

$$D_m = \sum_{i=1}^n \sum_j w_{j,i}^m \|x_j - x_{c,i}\|^2. \quad (14)$$

where  $x_j$  denotes a data point,  $x_{c,i}$  is the estimated center of the  $i$ th cluster,  $n$  denotes the number of clusters,  $m$  is a fuzzifier number larger than 1 and  $w_{j,i}$  is a weight that represents the degree of point  $x_j$  belonging to the  $i$ th cluster, where

$$\sum_{i=1}^n w_{j,i} = 1. \quad (15)$$

The FCM partitioning iteration starts with assigning the initial fuzzy weighting values  $w_{j,i}$ , and the assigned values should satisfy equation (15). It attempts to minimize  $D_m$  in equation (14), and the corresponding centers of the clusters and the fuzzy weighting value  $w_{j,i}$  are updated obeying the following two equations:

$$x_{c,i} = \frac{\sum_j w_{j,i}^m x_j}{\sum_j w_{j,i}^m}. \quad (16)$$

$$w_{j,i} = \frac{1}{\sum_{i=1}^n \frac{1}{\|x_j - x_{c,i}\|^{\frac{2}{m-1}}}}. \quad (17)$$

The sum of the weights of a point belonging to all the  $n$  clusters is constrained to 1. Equation (17) implies that the point having the longer distance from the cluster center has the smaller value of the weight contributing to the center. The iteration repeats until the  $D_m$  in equation (14) converges to a small value by means of the previous center updating and weight computing. As a result, a point is assigned as a member of the cluster having the highest weighting value  $w$ .

#### 3.3 K-Means Algorithm

K-means algorithm is used to segment a group of points into  $n$  clusters as FCM, and each feature point is consequently assigned to the cluster with the nearest center. However, unlike FCM algorithm, data points classified by means of the k-means algorithm are assigned to only a cluster with the binary degree simply either 1 or 0 [29][30]. Therefore, the variable  $w$  in equation (14) mentioned in FCM to evaluate the degree is abandoned in the k-means algorithm. The idea of k-means algorithm is to minimize the sum of the square distance

(SSD) within clusters and can be demonstrated by the function as below:

$$\arg \min_C \sum_{i=1}^n \sum_{x_j \in C_i} \|x_j - x_{c,i}\|^2. \quad (18)$$

where  $x_{c,i}$  is the center of the  $i$ th cluster  $C_i$  derived from the equation:

$$x_{c,i} = \frac{1}{N} \sum_{x \in C_i} x_j \quad (19)$$

where  $N$  is the number of points belonging to the  $i$ th cluster  $C_i$ .

In application, k-means algorithm begins with  $n$  points chosen randomly as the centers of the clusters. The distance from each point to the cluster centers is computed iteratively and every point is observed to be assigned to the cluster with the center closest to it. Next, the centers of the new clusters are also calculated by equation (19), and the new clusters and the distance from each point to the centers of clusters are repeatedly updated until the centers of the clusters no longer change.

#### 4. Experiment and Results

In this section, 45 images selected from 15 CT dataset with parotid pathology are used in experiments to evaluate the method applied to segment the soft tissues. The CT dataset are gathered from Cathay General Hospital and selected by the clinical experts. The CT images are scanned by Phillip Brilliance 64 scanner and the pixel spacing of the CT images is  $0.78 \times 0.78 \text{mm}^2$ . The regions of interest of a slice are selected at both sides of the cervical spine and the air path near the central axis of the body. Morphological methods including as erosion, dilation and labeling algorithms are introduced to extract the regions after segmentation. The performance was evaluated by comparing the segmentation results with the overlap fraction of lesions outlined by the

clinical experts. Detection rate (D.R.) defined as the equation (20) below is used to evaluated the result.

$$D.R. = 1 - Err \quad (20)$$

$$Err = \frac{A_{wrong}}{A_{whole}} \quad (21)$$

where  $A_{wrong}$  denotes the area of incorrect classification and  $A_{whole}$  denotes the area of the segmented region covering the target.

The features are derived using the Gaussian filter with the length of 3. The one to two levels decomposition appears more reasonable result than the decomposition in higher levels in the image quality. The three segmentation algorithms mentioned above are compared and some of the results are demonstrated. The pixels in those images are labeled with colors to represent the classes of the segmentation.

Figure 6 is the parotid region in an H&N CT with the lesion shown in weak boundaries between the adjacent normal tissues, and the segmentation results using different methods are compared and presented in Figure 7 to 10. Figure 7.(a) to Figure 7.(d) illustrate the segmentation results using the mean-shift algorithm with different bandwidth. It can be seen that the bandwidth seriously affects the segmentation result. But the appropriate bandwidth depends on the image qualities and is different in each case. Therefore, it is hard to be decided automatically. Segmentation with small bandwidth may result in over-segmentation as shown in Figure 7.(a), but points may be wrongly gathered into large groups if the bandwidth goes too large. Figure 8.(a) to Figure 8.(e) show the results using FCM with different fuzzifier number  $m$  indicated in equation (14). It can be seen that the segmentation performs better with the smaller  $m$  which performs in harder clustering. Figure 9.(a) to Figure 9.(d) show the results using k-means algorithms applying different image features. Features involving detail information result in more efficient segmentation than only scaling information, especially at the boundaries between tissues. Figure 10 is another test parotid image with the lesions of higher contrast at the boundaries. The performance of segmentation using different methods is also compared in the experiments and the results are shown in Figure 11 to Figure 13.

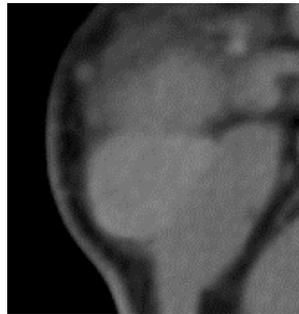
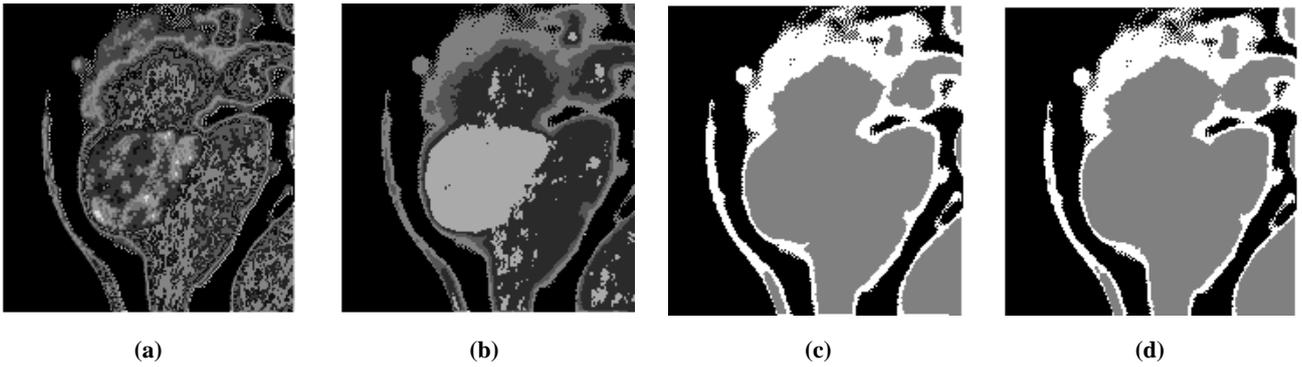
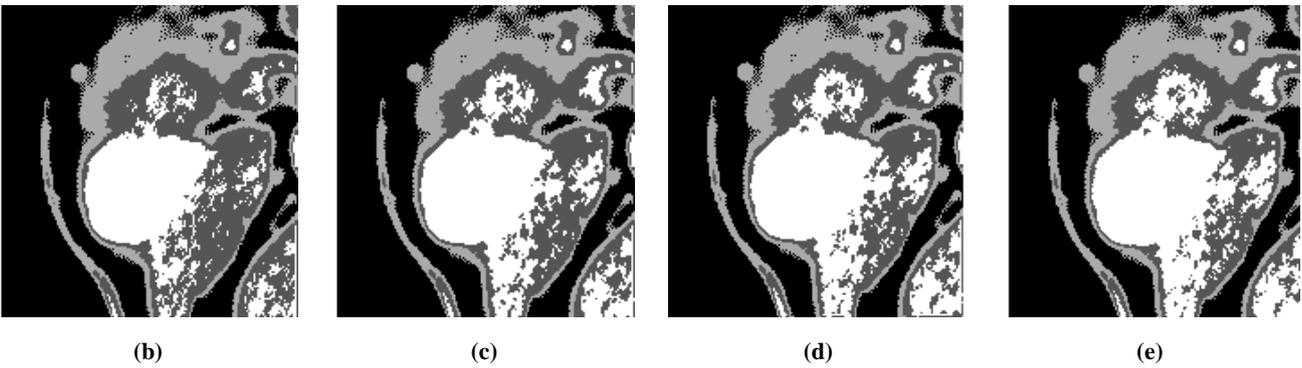
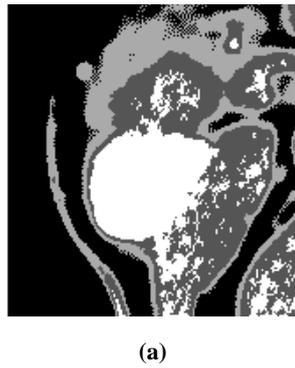


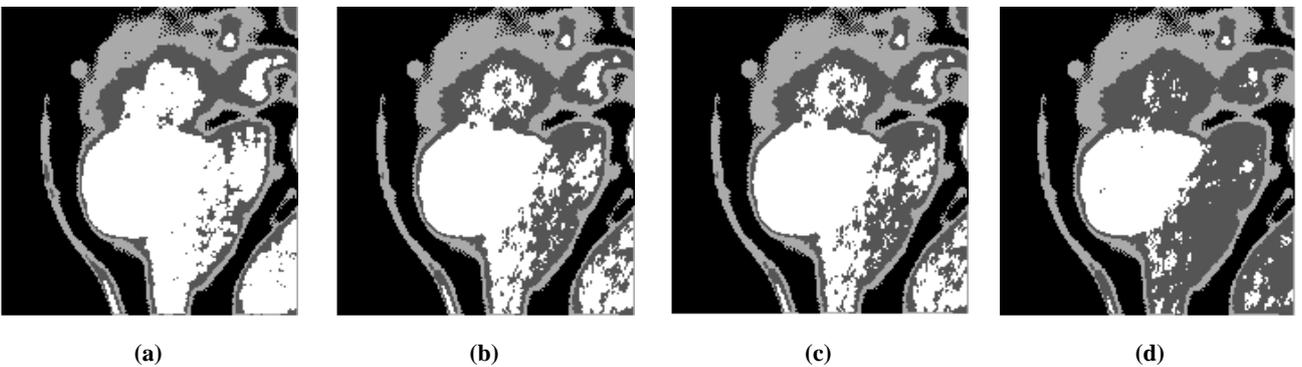
Figure 6: The ROI of the parotid in a CT slice



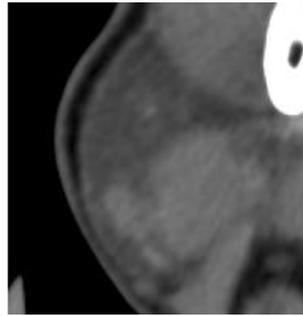
**Figure 7:** Segmentation results of Figure 6 using the mean-shift method with different bandwidth (a) bandwidth=25 (b) bandwidth=50 (c) bandwidth =75 (d) bandwidth =100.



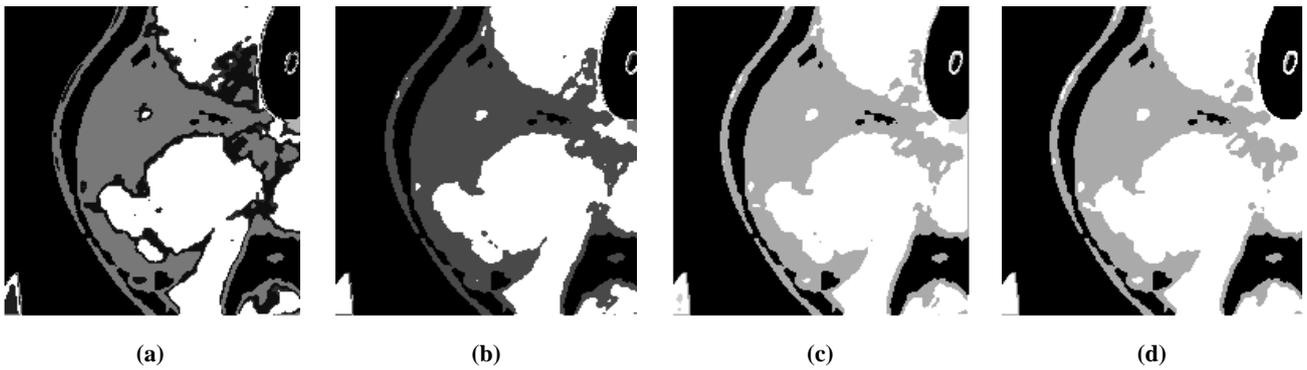
**Figure 8:** Segmentation results of Figure 6 using FCM with different fuzzier number  $m$  (a)  $m=2$  (b)  $m=5$  (c)  $m=10$  (d)  $m=15$  (e)  $m=20$



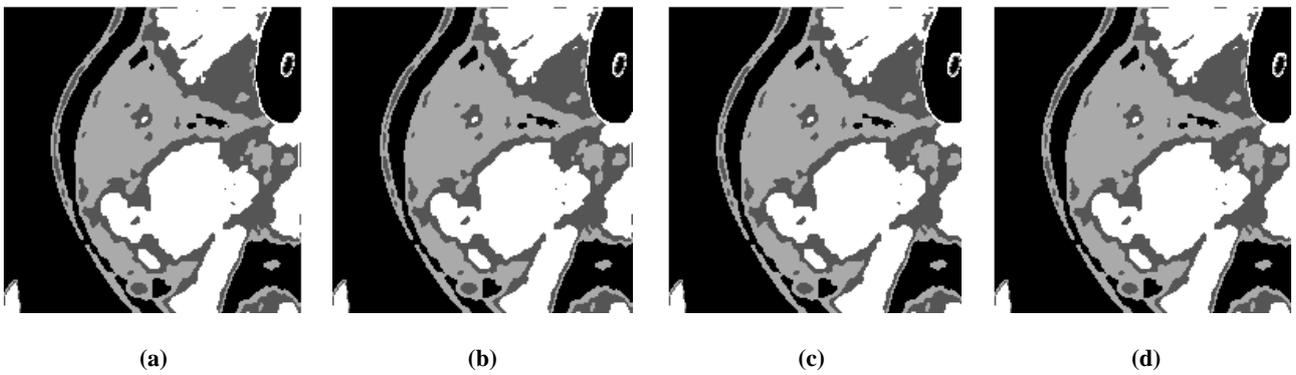
**Figure 9:** Segmentation results of Figure 6 using the k-means method (a) using only the gray level feature (b)(c)(d) the weight of the gray-level feature from high (b) to low (d).



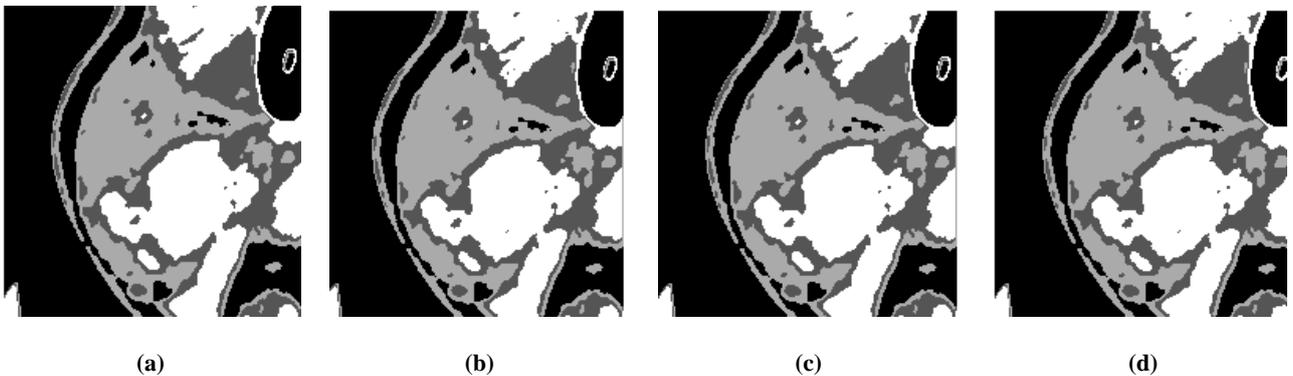
**Figure 10: The ROI of the parotid in a CT slice**



**Figure 11: Segmentation results of Figure 10 using with the mean-shift method using different bandwidth (a) bandwidth=25 (b) bandwidth=50 (c) bandwidth =75 (d) bandwidth =100.**



**Figure 12: Segmentation results of Figure 10 using with FCM method using different m (a) m=5 (b) m=10 (c) m=15 (d) m=20.**



**Figure 13: Segmentation results of Figure 10 using the k-means method (a) using only the gray level feature (b)(c)(d) the weight of the gray-level feature from high (b) to low (d).**

## 5. Discussion and Conclusions

Segmentation using mean-shift algorithm can perform satisfying results if an appropriate bandwidth is chosen for clustering computation. However, the experiments indicate that the appropriate bandwidth in different cases varies in a large range and need to be decided manually. The performance of FCM is inferior to the k-means algorithm shown in the experiments, especially for the weak boundaries in the images. K-means algorithm performs in relatively high D.R in those experiments, and the segmentation applying more features about the detail information including the variation energy, entropy and variance performs better than the segmentation applying only the gray-level distribution or the scaling information derived from the approximation coefficients in DWT. The average detection rate utilizing the same features is listed in Table 1.

**Table 1: Average detection rate with different segmentation methods.**

Method	Avg. Rate %
Mean-shift Algorithm	91.151
FCM Algorithm	84.277
K-means Algorithm	95.314

Conclusively, the distinctive feature descriptors used for textures are applicable in the segmentation for the soft tissues of parotid glands and the lesions. Besides, hard clustering is performing better than the soft clustering in this work.

In this paper, we proposed a method for segmenting the soft tissues of the parotids and lesions in the CT images with pathology. The image features derived from wavelet analysis without decimation are involved to improve the performance of segmentation utilizing only gray level usually applied in CT images. Unsupervised clustering algorithms are utilized in this work, and the segmentation results in the experiment demonstrate that the lesion regions can be effectively separated from the adjacent tissues with the distinctive features. If the scale of analysis gets higher, the boundaries of the segmentation may be more blurred such that some lesions with small size may be easily neglected. The delineation of the extracted regions can be helpful for evaluating the initial condition of the active contour, and the segmentation of the regions of the skeptical lesions is beneficial for vision-aided assisting diagnosis. Feature work will involve the combination of the segmentation methods with shape analysis and recognition methods for medical automation approaches.

## 6. REFERENCES

- [1] Duay, V., Houhou, N., Bach Cuadra, M., Schick, U., Becker, M., Allal, A.S. and Thiran, J. P. 2009. Segmentation of head and neck Lymph node regions for radiotherapy planning using active contour-based atlas registration. *IEEE Journal of selected topics in signal processing*, Vol. 3, No.1 (Feb. 2009), 135-147.
- [2] Teng, C. C., Shapiro, L. G. and Kalet, I. 2006. Automatic Segmentation of Neck CT Images. *Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems*.
- [3] Xu, X.Q., Lee, D. J., Antani, S. and Long, L. R. 2004. *Proceedings of the 17th IEEE Symposium on Computer-based Medical Systems*.
- [4] Feulner1, J., Zhou. S. K. , Hammon, M., Hornegger, J. and Comaniciu, D. 2011. Segmentation-based features for lymph Node Detection from 3-D Chest CT. *Proceeding MLMI'11 Proceedings of the Second international conference on Machine learning in medical imaging*
- [5] Barbu, A., Suehling, M., Xun, X., Liu, D., Zhou, S. K. and Comaniciu, D. 2012. Automatic Detection and Segmentation of Lymph Nodes From CT Data. *IEEE Transactions of Medical Imaging*, Vol. 31, No. 2(Feb. 2012), 240-250.
- [6] Tara, A. W., Henk, P. B., Henriëtte E. W. and Johannes, A. L. 2009. Delineation guidelines for organs at risk involved in radiation –induced salivary dysfunction and xerostomia, Vol. 93, No. 3(Dec. 2009), 545-552.
- [7] Duay, V., Houhou, N. and Thiran, J. P. 2005. Atlas-based segmentation of medical images locally constrained by level sets. *ICIP 2005. IEEE International Conference on Image Processing*.
- [8] Liliane, R. and Grégoire, M. 2010. Multi-atlas Based Segmentation: Application to the Head and Neck Region for Radiotherapy Planning. *MICCAI Workshop Medical Image Analysis for the Clinic - A Grand Challenge*.
- [9] Gorthi, S., Bach C. M., Schick, U., Tercier, P. A., Allal, A. S. and Thiran, J. P. 2010. Multi-Atlas based Segmentation of Head and Neck CT Images using Active Contour Framework. *MICCAI workshop on 3D Segmentation Challenge for Clinical Applications*.
- [10] Yang, J., Zhang, Y., Zhang, L. and Dong, L. 2010. Automatic Segmentation of Parotids from CT Scans Using Multiple Atlases. *Medical Image Analysis for the Clinic - A Grand Challenge*.
- [11] Commowick, O. and Malandain, G. 2007. Efficient selection of the most similar image in a database for critical structures segmentation. *Medical Image Computing and Computer Assisted Intervention*.
- [12] Lauric, R. and Frisken, S. 2007. Soft segmentation of CT brain data. *Technical Report TR-2007-3 Tufts University, MA, Tech. Rep.*
- [13] Asmita, A. M., Singhai, J. and Shrivastava, S. C. 2011. Automatic Threshold based Liver Lesion Segmentation in Abdominal 2D-CT Images. *International Journal of Image Processing*, Vol. 5, No. 2(Sep. 2011), 166-176.
- [14] Semler, L., Dettori, L. and Furst, J. 2005. Wavelet-based texture classification of tissues in computed tomography.

- Proceedings 18th IEEE Symposium on Computer-Based Medical Systems.
- [15] Latif, G., Kazmi, S. B., Jaffar, M. A. and Anwar, M. M. 2010. Classification and Segmentation of Brain Tumor using Texture Analysis. International Conference on Recent advances in artificial Intelligence, Knowledge Engineering and Databases.
- [16] Kruggel, F., Paul, J. S. and Gertz, H. J. 2007. Texture-based segmentation of diffuse lesions of the brain's white matter. *NeuroImage*, Vol. 39(Oct. 2007), 987-996.
- [17] Livens, S., Scheunders, P., Wouwer, G. and Dyck, D. 1997. Wavelets for texture analysis, an overview. International Conference on Image Processing and Its Applications.
- [18] Arivazhagan, S. and Ganesan, L. 2003. Texture classification using wavelet transform. *Pattern Recognition Letters*, Vol. 24, No. 9(June 2003), 1513-1521.
- [19] Fan, G. and Xia, X. G. 2003. Wavelet-based texture analysis and synthesis using hidden Markov models. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 50, No. 1(Jan. 2003), 106-120.
- [20] Bayram, K. and Watsuji, N. 2003. Using Wavelets For Texture Classification. *IJCI Proceedings of International Conference on Signal Processing*, Vol. 1, No. 2(Sep. 2003), 920-924.
- [21] Wouwer, G., Scheunders P. and Dyck, D. 1999. Statistical texture characterization from discrete wavelet representations. *IEEE Transactions on Image Processing*, Vol. 8, No. 4(Apr. 1999), 592-598.
- [22] Laine, A. and Fan, J. 1993. Texture classification by wavelet packet signatures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 11(Nov. 1993), 1186-1191.
- [23] Bashar, M.K., Matsumoto, T. and Ohnishi, N. 2003. Wavelet transform-based locally orderless image for texture segmentation, *Pattern Recognition Letters*, Vol.24, No. 15, (Nov. 2003), 2633-2650.
- [24] Holschneider, M., Kronland-Martinet, R., Morlet, J. and Tchamitchian, P. 1989. A real-time algorithm for signal analysis with the help of the wavelet transform, In *Wavelets, Time-Frequency Methods and Phase Space*, 289-297.
- [25] Cheng, Y. 1995. Mean Shift, Mode Seeking, and Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8(Aug. 1995), 790-799.
- [26] Comaniciu, D. and Meer, P. 2002. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5(May, 2002), 603-619.
- [27] Nock, R. and Nielsen, F. 2006. On weighting clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 8(Aug. 2006), 1223-1235.
- [28] Ahmed, M., Yamany, S. M., Mohamed, N., Farag, A. A. and Moriarty, T. N. 2002. A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data, Vol. 21, No. 3(Mar. 2002), 193-199.
- [29] Hartigan, J. A. 1975. *Clustering Algorithms*. John Wiley & Sons Inc. ISBN-13: 978-0471356455.
- [30] Renato, C. A. and Boris, M. 2012. Metric, feature weighting and anomalous cluster initializing in K-Means clustering, Vol. 45, No. 3(Mar. 2012), 1061-1075.