FFT based interpolation and 3D level set methods for 3D throat region segmentation from MRI data*

Sean Campbell, Trushali Doshi, John Soraghan *Senior Member, IEEE,* Lykourgos Petropoulakis, Gaetano Di Caterina, Derek Grose, Kenneth MacKenzie

***Abstract*— A new algorithm for 3D throat region segmentation from magnetic resonance imaging (MRI) is presented. The proposed algorithm initially pre-processes the MRI data to increase the contrast between the throat region and its surrounding tissues and also to reduce artifacts. Isotropic 3D volume is reconstructed using Fast Fourier Transform based interpolation. Furthermore, a cube encompassing the throat region is evolved using level set method to form a smooth 3D boundary of the throat region. The results of the proposed algorithm on real and synthetic MRI data are used to validate the robustness and accuracy of the algorithm.**

I. INTRODUCTION

Segmentation of the throat region, commonly known as the pharynx or upper airway region, from magnetic resonance imaging (MRI) is important in many applications. The detected throat region can, for instance, be used as additional information for automatic segmentation of oropharynx tumors [1]. Pharynx segmentation from MRI can additionally assist in the study of soft tissue structures of the pharynx for sleep related disorders [2-3] or for the detection of pharynx stenosis. For surgical planning in spine surgery, the segmentation of pharynx region is a pre-requisite [4].

For pharynx detection, a few approaches have been reported in the literature. In [1], a Hessian matrix of low intensity pixels is analyzed for the pixels of the throat region. The region with the largest repeated eigenvalues is selected as the throat region. A semi-automatic approach utilizing a set of general rules is applied for pharynx extraction from MRI in [2]. The techniques in [1-2] extract pharynx regions from 2D MRI slices and thus require further processing to reconstruct a volume in 3D. Another semi-automatic approach using fuzzy connectedness for segmentation of the upper airway lumen is presented in [3]. In this approach, segmentation results need to be verified by the operator as results may include extraneous air-space regions. The trachea or the lower airway region has the same characteristics as pharynx on axial MRI images. Researchers in [4] utilize initial landmarks to delineate the trachea in 2D and used that as an initialization for active contours to extract the structure in 3D. However, this technique is applied in the sagittal plane. Active contour [5] segmentation of the trachea using natural cubic splines was presented in [6]. Initialization of a snake contour using a threshold based Canny edge filter permit this technique [6] only for computed tomography (CT) images. In a semi-automatic approach [7], a cost function is created from the gradient and knowledge of trachea. A graph based approach is utilized to find the least cost surface which is labeled as the trachea. A fully automatic approach using intensity clustering and graph cut technique is proposed in [8]. These techniques [7-8], make a few assumptions about roundness (circularity) and area of the trachea which cannot be assumed for the throat regions in this work.

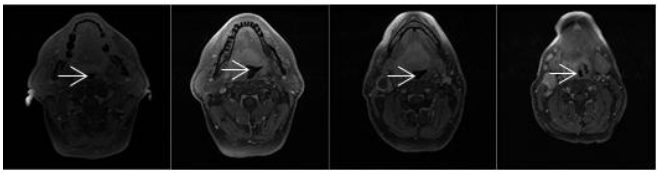


Fig.1: Real MRI data demonstrating geometric variability of throat regions (white arrow) and low contrast.

Level set methods (LSM) [9-10] have been employed in the past for segmentation of the trachea [11] or trachea like structures such as abdominal aortic aneurysms [12] from CT imaging. LSM is a numerical technique which deforms the curve C implicitly by evolution of a higher dimensional function ϕ. This implicit representation allows change of topology such as splitting and merging and is useful for shape recovery of complex anatomical structures. The evolution equation of ϕ is given as

$$\frac{\partial \phi}{\partial t} = F|\nabla \phi| \tag{1}$$

where F is speed function that generally incorporates internal and external energy functions. An internal energy function serves as a smoothness regulator whereas external energy functions formed from image information drive ϕ to desired object boundaries. LSM generally require a regular cube mesh for evolution [12]. However, the voxels from MRI data in this work are not cubic voxels (z -dimension of voxel is significantly larger than x - and y -dimensions). Interpolation of the MRI volume in the z -direction is performed using Fourier Interpolation (FT), which effectively employs the Discrete Fourier Transform (DFT) or Fast Fourier Transform (FFT) to produce cubic (isotropic) voxels. This technique was chosen for interpolation due to its accuracy and simplicity in comparison to other interpolation techniques [13].

*Research supported by Beatson Cancer Charity.

Sean Campbell is with the Dept. of Electronic and Electrical Engineering, University of Strathcylde, Glasgow G1 1XW, UK. (email: sean.campbell.100.2013@uni.strath.ac.uk)

Trushali Doshi (email: trushali.doshi@strath.ac.uk), John Soraghan, Lykourgos Petropoulakis and Gaetano Di Caterina are with the Centre for excellence in Signal and Image Processing, Dept. of Electronic and Electrical Engineering, University of Strathcylde, Glasgow G1 1XW, UK.

Derek Grose is with the Beatson West of Scotland Cancer Centre, Gartnavel Hospital, Glasgow G12 0YN.

Kenneth MacKenzie is with the Glasgow Royal Infirmary, Glasgow, G4 0SF.

Segmentation of 3D throat regions in MRI is a challenging task due to geometry variability, the merging of both airways, the splitting of the hypopharynx into the trachea and the esophagus, the potential presence of cancerous regions in addition to low contrast and inherent MRI artifacts. Typical examples of MRI slices used in this work are illustrated in Fig. 1. Low contrast and geometric variability of the throat region is clearly visible in Fig.1. This paper presents a robust technique which successfully overcomes the above challenges to obtain smooth 3D model structure of the throat region from contrast-enhanced T1-weighted MR images. In this work segmentation is formulated as a continuous boundary detection problem. This work utilizes the high contrast existing between the intensity of the throat region and that of the surrounding tissues for the evolution of level set to segment the throat region.

The organization of the remainder of the paper is as follows: Section II describes the proposed algorithm for 3D throat region segmentation. Section III demonstrates the results on a synthetic and a real MRI dataset. Section IV presents discussion of this work and some future work, including potential improvements.

II. 3D THROAT REGION SEGMENTATION

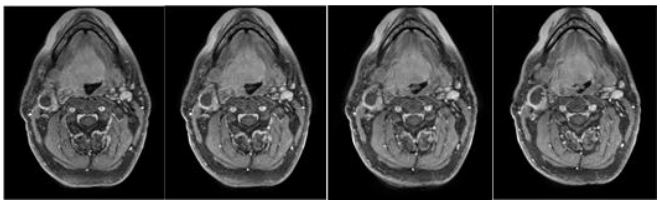
The proposed 3D throat region segmentation process comprises three steps: a) pre-processing, b) FFT based interpolation and c) LSM segmentation of the throat region.

A. Pre-processing

Pre-processing is applied separately to each axial (2D) MRI slice to reduce the data size and enhance the contrast between throat region and its neighboring tissues. A series of morphological operations are applied in conjunction with edge detection techniques for the removal of a significant proportion of background noise, while preserving the primary features of the images. After noise removal, the overall size of the volume is deduced through detecting the first non-zero intensity value along each dimension and then performing cropping. In order to increase the contrast (intensity difference) a background brightness preserving contrast enhancement (BBPHE) technique [14] is applied. This technique improves contrast while preserving throat region brightness. Finally, intensity inhomogeneity (IIH) [15] observed in the MRI data used in this work is reduced using the technique described in [15]. A few examples of MRI slices after the pre-processing step are illustrated in Fig. 2(a) and Fig 2(d).

B. FFT based interpolation

The voxel of pre-processed MRI data in this work is non-isotropic with a size 0.47x0.47x3.5mm^3. In order to generate isotropic voxels, interpolation is carried out using Fourier Interpolation [16] in the z-direction only. In this interpolation technique, the original data is transformed to Fourier domain using the FFT. Zeros are appended to center of the data, with the Nyquist criterion being satisfied in the case of an even sized dimension, and the inverse FFT is performed. The interpolation technique can be represented using



(a) (b) (c) (d)

Fig.2: Pre-processed and interpolated MRI slices: (a) and (d) pre-processed real MRI slices and (b) and (c) interpolated slices using FFT interpolation

$$Y(k)=\begin{cases} X(k), & for\ \ k=0,\ldots,\frac{N}{2}-1 \\ X(\frac{N}{2})/2, & for\ \ k=\frac{N}{2},NL-\frac{N}{2} \\ 0, & for\ \ k=\frac{N}{2}+1,\ldots,NL-\frac{N}{2}-1 \\ X(K+N-NL), & for\ \ k=NL-\frac{N}{2}+1,\ldots,NL-1 \end{cases} \quad (2)$$

where $X(k)$ is the FFT of the input signal $x(n)$ of length N and $Y(k)$ is the zero-padded FFT of $X(k)$ with length NL, L denoting the interpolation factor. The interpolated signal $y(n)$ is obtained by taking the NL-point inverse FFT of $Y(k)$. The $1/L$ amplitude loss made by interpolation is compensated by multiplying signal $y(n)$ by L [16]. Defining Y(k) as in Eq. (2) maintains the complex-conjugate circular symmetry which characterizes the FFT of real-valued signals.

Fig. 2 shows two interpolated slices (Fig. 2(b) and Fig.2 (c)) obtained using the technique between two consecutive real MRI slices (Fig. 2(a) and Fig. 2(d)). The interpolated slices are similar to real MRI slices and demonstrate minimal error. The decrease in voxel size in the z-direction due to these interpolated slices increases the accuracy of the level set segmentation and helps to obtain a smooth 3D throat region. The volume is reconstructed in 3D using both the real and interpolated slices.

C. Throat volume segmentation with level sets

For throat region segmentation from the reconstructed volume, a cube is defined in an arbitrary location encompassing the throat region and evolved using a level set function. In this work, the speed function F (1) used for level set evolution is given as [17]

$$F=\alpha(\varepsilon-|I(x,y,z)-T|)+(1-\alpha)\nabla.\frac{\nabla\phi}{|\nabla\phi|} \quad (3)$$

where $\nabla.\frac{\nabla\phi}{|\nabla\phi|}$ is the mean curvature of the surface, serving as the internal energy function and preserving smoothness. $\alpha\in[0\ \ 1]$ is a free parameter which controls the weighting between data term (external energy) and smoothing term. In this work a signed distance function, based on Euclidean distance, is utilized as the level set function ϕ and a simple data term depending solely on the input intensity I at a point in space is considered. This data term is suitably simple because there is high intensity contrast between the throat region and surrounding tissues. In (3), T controls the intensity of the region to be segmented and ε controls the intensity (gray level) range around T that could be

considered inside the object. A narrow band method is employed in our implementation of level set for computational efficiency. This LSM robustly handles splitting and merging and obtains a smooth 3D structure of the throat region. The isolated noisy voxels segmented by the LSM technique are reduced only by considering the largest segmented volume as the throat region.

III. Experimental Results

A. Synthetic MRI dataset

To obtain quantitative results for the proposed algorithm, a simple synthetic dataset was generated. Synthetic MRI slices were reconstructed using the Shepp-Logan head phantom [18]. The throat region was constructed as low intensity ellipses of varying sizes, approximating the shape of a conical sphere. In total 32 slices were reconstructed with varying size of the throat region. The in-plane resolution was 256x256 and the number of slices was interpolated to equal this, essentially creating a 256x256x256 volume.

The initial mask of LSM covered a total volume of 96x96x256, approximately 14% of the total volume, and included a portion of all features of the model. Fig. 3(a) illustrates a number of individual slices with the throat section highlighted and, despite some small rippling at the edges of each ellipse; the segmented volume (Fig. (3(b)) is predominantly in-line with expectations. The model employed a low magnitude for α for smoothness but instead of issuing a strict number of iterations for the LSM to be performed, a stopping criterion was given. Effectively, this criterion checked the number of differences between the narrow-bands of the previous zero-order level set and the current, prematurely cutting off the algorithm if the number fell below a threshold, indicating convergence.

For quantitative results, the ground truth for throat region was obtained by applying the threshold and selecting the smallest low intensity region. The throat region segmented using the LSM were compared with a ground truth on a slice by slice fashion. The true positive (TP), true negative (TN), false positive (FP) and false negative (FN) were calculated for 16 synthetic MRI slices. The F-measure [0 1] which estimates algorithmic accuracy was calculated as:

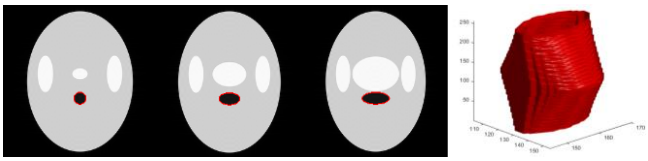
$$precision = \frac{TP}{TP + FP}, recall = \frac{TP}{TP + FN}$$

$$F - measure = 2.\frac{precision.recall}{precision + recall} \quad (4)$$

Fig. 4 shows F-measure value for 16 slices. It can be observed that F-measure value is >0.99 for all slices. However, low value of F-measure (lowest 0.9936) can be attributed to FN values (max 6 pixels) as the proposed technique tends to underestimate the throat region in some MRI slices. This underestimation is primarily due to the rippling that occurs at the edges of ellipse. However, an increase in resolution will decrease the rippling effect and increase the algorithm accuracy, at the expense of additional computational time.

B. Real MRI Dataset

For the real dataset, contrast-enhance T1-weighted MRI scans were obtained using a GE Medical Systems Signa HDxt 1.5-Tesla scanner. 11 real MRI slices with voxel resolution of 0.47x0.47x3.5mm^3 were selected to validate the



(a) (b)

Fig. 3: Synthetic MRI dataset (a) throat region boundary (red outline) from 3 slices (b) segmented throat volume from 256x256x256.

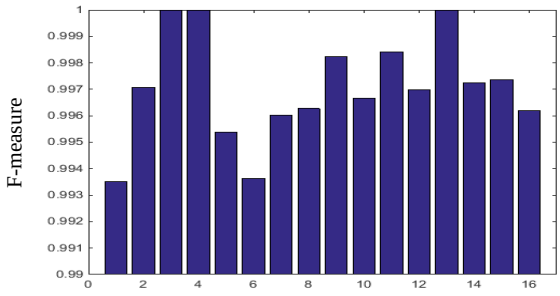


Fig. 4: Quantitative analysis of the proposed algorithm from a synthetic dataset.

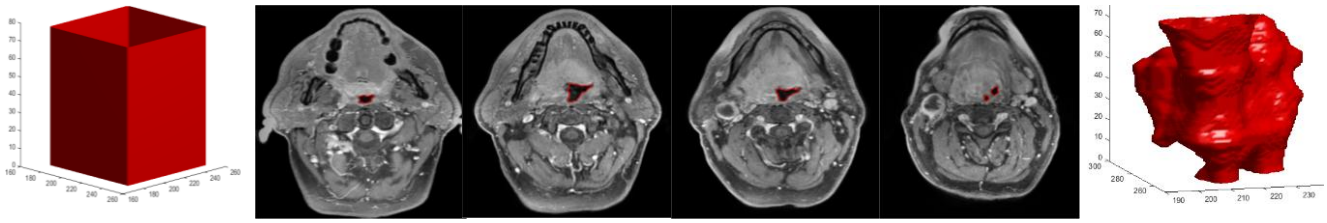
algorithm. To make voxels approximately isotropic, 7 slices were interpolated in between consecutive real slices. A total of 81 (11 real + 70 interpolated) MRI slices were used in this work. The voxel resolution after interpolation step was 0.47x0.47x0.5mm^3. The LSM algorithm was verified on the 3D volume reconstructed using these 81 slices.

For initialization of the LSM, a simple mask encompassing the throat region with a volume of 78x78x78, corresponding to approximately 10% of the total volume, was used and is demonstrated with Fig.5 (a). Fig. 5(b) illustrates a number of slices with the throat section highlighted, with the individual segments being visibly accurate. In addition, Fig. 5(c) demonstrates the final segmented volume. From Fig. 5(c) it is clear that the segmented volume is smooth and no further processing is required. The LSM algorithm which produced this segmentation employed a low magnitude for α ensuring a smooth result but a slightly higher number of iterations for convergence.

IV. Discussion

In this paper we proposed a new method for volumetrically segmenting the throat region from a number of MRI slices. Our model performs the segmentation in 3D, eliminating the need for post-processing reconstruction and smoothing algorithms if one were to instead perform the segmentation on a slice by slice basis. The pre-processing step implemented in this work allows reduction of false positives. FFT based interpolation helped to increase the resolution along z-direction, thus increasing the accuracy in throat region segmentation.

The model discussed in this paper possesses a relatively simplistic data function in LSM but the results illustrate it is sufficient for performing the segmentation process for relatively high contrasting regions. By coupling the data and smoothing weights in (3), one simplifies the selection process. The weights were decoupled, for exploratory purposes, and through multi-variable optimization techniques, with optimization based upon differences between obtained segmentations and their ground truths, subtle differences in the weights led to equally subtle superior



(a) (b) (c)

Fig. 5: Real MRI dataset (a) initialization for LSM (cube of size78x78x78) (b) representative 2D slices with throat region segmentation results (red outline) (c) final 3D structure of the throat region.

results.

The main advantages of the proposed method is that is makes no assumption regarding the individual shapes or areas. Furthermore using the method it becomes unnecessary to smooth the initial volume with Gaussian filters prior to segmentation, offering some small computational saving and preserving data fidelity. The algorithm itself is generally practicable in a number of areas and can be employed to segment any object with good contrast between desired object and its neighbors.

There are several areas of the overall process which still warrant development. Computationally, the narrow-band LSM algorithm has been constructed in a sequential manner and a massive performance boost could undoubtedly be attained if the procedure were carried out in parallel, potentially allowing real-time tuning of the parameter values indicated in (3). Alternatively, one could opt for a different route and investigate automatic initialization based upon a priori information on desired region(s) of interest or learned through training data, though naturally this will limit the application domain. The current model still requires thorough testing with large and varied datasets, particularly in cases where models exhibit weak-edges to investigate how successfully the data function performs, but it certainly exhibits promise.

ACKNOWLEDGMENT

The authors would like to acknowledge Beatson Cancer Charity for their financial support with this study.

REFERENCES

[1] T. Doshi, J. Soraghan, L. Petropoulakis, D. Grose and K MacKenzie, “Modified Fuzzy C-means Clustering for Extraction of Tongue Base Tumor from MRI data”, in Proc. of IEEE EUSIPCO, Lisbon, Portugal, 2014, pp. 2460-2464.

[2] T. Ivanovska, J. Dober, R. Laqua, K. Hegenscheid, and H. Volzke, “Pharynx Segmentation from MRI Data for Analysis of Sleep Related Disoders” in Proc. of International Symposium on Visual Computing, Greece, 2013, pp. 20–29.

[3] J. Liu, J. Udupa, D. Odhnera, J. McDonough and R. Arens, “System for Upper Airway Segmentation and Measurement with MR Imaging and Fuzzy Connectedness”, Academic Radiology, vol. 10, no. 1, pp. 13-24, Jan 2003.

[4] S. Seifert, I. Wachter, G. Schmelzle, and R. Dillmann, “A Knowledge-Based Approach to Soft Tissue Reconstruction of the Cervical Spine”, IEEE Trans. Medical Imaging, vol. 28, no. 4, pp. 494-507, 2009.

[5] M. Kass, A Witkin, and D. Terzopoulos, “Snakes: Active contour models.” in Proc. 1st International Conf. Computer Vision, 1987, pp. 259- 268.

[6] R. Valdes, O. Yanez-Suarez, and V. Medina. “Trachea Segmentation in CT Images using Active Contours”, in Proc. of IEEE EMBC, Chicago, USA, 2000, vol. 4, pp. 3184 – 3187.

[7] R. Amendola, J. Reinhardt, Y. Sata, M. Zimmerman, H. Diggelmann and D Kacmarynski, “Graph based segmentation of the pediatric trachea in MR images to model growth”, in Proc. SPIE Medical Imaging, Florida, USA, 2013.

[8] T. Ivanovska, E. Buttket, R. Laqua, H. Volzke and A. Beulet, “Automatic Trachea Segmentation and Evaluation from MRI Data Using Intensity Pre-Clustering and Graph Cuts”, in Proc. of ISPA, Dubrovnik, Croatia, 2011, pp. 513-518.

[9] J. Sethian. Level Set Methods and Fast Marching Methods. Cambridge University Press, 1999.

[10] S. Osher, R.Fedkiw. Level Set Methods and Dynamic Implicit Surfaces. Springer-Verlag, 2003.

[11] T. Deschamps, P. Schwartz, and D. Trebotich, “Air-flow simulation in realistic models of the trachea”, in Proc. of IEEE EMBC, San Francisco, CA, 2004, pp. 3933 - 3936.

[12] D. Mageea, A. Bulpitt and E. Berry, “Level Set Methods for the 3D Segmentation of CT Images of Abdominal Aortic Aneurysms”, in Proc. Med. Image Understand. Anal., 2001, pp. 141–144.

[13] W. Hawkins, “FFT interpolation for arbitrary factors: a comparison to cubic spline interpolation and linear interpolation” in Proc. Nuclear Science Symposium and Medical Imaging Conference, Norfolk, VA, 1994, vol.3, pp. 1433 – 1437.

[14] T. Tan, K. Sim and C. Tso, “Image enhancement using background brightness preserving histogram equalisation”, IEEE Electronics Letters, vol. 48, no. 3, Feb 2012.

[15] O. Salvado, C. Hillenbrand, S. Zhang, and D. Wilson, “Method to Correct Intensity Inhomogeneity in MR Images for Atherosclerosis Characterization”, IEEE Trans Medical Imaging, vol. 25, no. 5, pp. 539-552, May 2006.

[16] J. Adams, “A subsequence approach to interpolation using the FFT,” IEEE Trans Circuits and Systems, vol. 34, no. 5, 1987, pp. 568 – 570.

[17] A. Lefohn, J. Kniss, C. Hansen, and R. Whitaker, “A Streaming Narrow-Band Algorithm: Interactive Computation and Visualization of Level Sets”, IEEE Trans. Visualization and Computer Graphics, vol. 10, no. 4, 2004, pp. 422-433.

[18] L. Shepp and B., Logan, "The Fourier Reconstruction of a Head Section", IEEE Trans. Nuclear Science, vol. 21, no. 3, Jun 1974, 21-43.