Scientific Research Publishing

Strategies for Segmenting the Upper Airway in Cone-Beam Computed Tomography (CBCT) Data

N. Kabaliuk[1], A. Nejati[2], C. Loch[3], D. Schwass[3], J. E. Cater[2], M. C. Jermy[1*]

[1]Department of Mechanical Engineering, University of Canterbury, Christchurch, New Zealand
[2]Department of Engineering Sciences, University of Auckland, Auckland, New Zealand
[3]Sir John Walsh Research Institute, Faculty of Dentistry, University of Otago, Otago, New Zealand
Email: mark.jermy@canterbury.ac.nz

Abstract

The wide availability, low radiation dose and short acquisition time of Cone-Beam CT (CBCT) scans make them an attractive source of data for compiling databases of anatomical structures. However CBCT has higher noise and lower contrast than helical slice CT, which makes segmentation more challenging and the optimal methods are not yet known. This paper evaluates several methods of segmenting airway geometries (nares, nasal cavities and pharynx) from typical dental quality head and neck CBCT data. The nasal cavity has narrow and intricate passages and is separated from the paranasal sinuses by thin walls, making it is susceptible to either over- or under-segmentation. The upper airway was split into two: the nasal cavity and the pharyngeal region (nasopharynx to larynx). Each part was segmented using global thresholding, multi-step level-set, and region competition methods (the latter using thresholding, clustering and classification initialisation and edge attraction techniques). The segmented 3D surfaces were evaluated against a reference manual segmentation using distance-, overlap- and volume-based metrics. Global thresholding, multi-step level-set, and region competition all gave satisfactory results for the lower part of the airway (nasopharynx to larynx). Edge attraction failed completely. A semi-automatic region-growing segmentation with multi-thresholding (or classification) initialization offered the best quality segmentation. With some minimal manual editing, it resulted in an accurate upper airway model, as judged by the similarity and volumetric indices, while being the least time consuming of the semi-automatic methods, and relying the least on the operator's expertise.

Keywords

Cone Beam CT, CBCT, Segmentation, Upper Airway, Nasal Cavity,

Pharyngeal Airway

1. Introduction

Respiratory illness affects a substantial number of people worldwide with the top five respiratory diseases accounting for 17.4% of all deaths and 13.3% of all Disability-Adjusted Life Years (DALYs) [1]. Breathing therapy devices are frequently used to treat breathing disorders such as chronic obstructive pulmonary disease (COPD) and obstructive sleep apnoea (OSA), and to assist premature babies as their lungs develop. Breathing therapies are also frequently administered to patients in critical or intensive care wards. The effectiveness of the therapy depends on the design of the device, the inspirational demand, face shape (mask or cannula fit) and airway shape (which determines flow resistance and gas mixing) [2]. Airway shape, the nasal cavity in particular, shows significant interpersonal variation [3] [4] [5] [6]. Knowledge of this variation is important for the design of better therapeutic devices, and requires segmentation of the airway from many individual scans.

In vivo measurements of pressure or CO_2 clearance are difficult due to the inaccessibility of the airway interior, and expensive to conduct with enough participants to be representative of the population. However gas mixing (CO_2 washout), flow and resistance can be studied using computational fluid dynamics (CFD) [7]-[18] or experimental methods using cast or 3D printed model airways [8] [19]-[24]. Accurate physical airway models are essential to these approaches. To make the models, or CFD meshes, shape data for both patient-specific and population-averaged airways are needed. Population-averaged airways have been generated by superimposing aligned individual patient-specific airway geometries [25], by averaging images of the airways slice-by-slice [26], by using used thin-plate spline deformations [27], or by applying a Fourier descriptor method to decompose airway shapes into morphological feature descriptors which can be averaged [12] [28]. Whichever method is followed, to generate airways representative of the population requires the segmentation of large numbers of scans, which is time consuming if done manually.

1.1. Data Sources for Human Upper Airway Studies

Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scans are valuable sources of anatomically correct 3D airway shape data. Segmentation is the process of extracting the structure of interest by identifying those image voxels which belong to the airway spaces based on their intensity, or patterns of intensity in a region of interest. Accurate segmentation depends on image quality and contrast, which vary with imaging modality.

MRI data offers the best tissue/air contrast, but movement artefacts may be present due to the long image acquisition time. In the upper airway, this particularly affects the tongue and soft palate.

Conventional helical multislice CT data is widely used as a source of accurate anatomical 3D models due to good spatial and contrast resolution, and limited movement artefacts, due to the short duration of the scan. It suffers from poor soft tissue differentiation, and streak and beam-hardening artefacts (dark streaks in the image between two high attenuation objects, *i.e.* metal, bone, dental fillings, or along the long axis of a single high attenuation object)which complicate the segmentation process, particularly around the oropharynx region [29] [30]. Each slice (2D image) is reconstructed independently in 2D and the slices are stacked to obtain a 3D reconstruction. The spatial resolution is usually determined by the distance between slices.

Cone-Beam CT or CBCT is a type of CT imaging widely used for various applications in dentistry, and hence is an abundant source of anatomical craniofacial image data. CBCT scanners have grown in popularity in clinics due to their lower cost and lower radiation dose compared to conventional CT [31]. A cone-shaped X-ray beam directed through the middle of the field of view onto a planar area detector [32] [33]. A complete set of CBCT images (from 150 to 600) is acquired in a single X-ray tube rotation, in a complete or partial arc. This reduces the scanning duration, movement artefacts and radiation exposure compared to conventional CT. Volumetric reconstruction is performed on acomplete set of raw 2D projected images. This allows for roughly cubic voxels *i.e.* equal spatial resolution in all directions. The spatial resolution is primarily dependant on the detector pixel size (typically 0.09 to 0.4 mm), and the spatial resolution is usually greater than conventional CT. The quality of the reconstructed shape data in CBCT is affected by the number of projections taken. A greater number of projections takes longer and increases radiation dose, but improves spatial resolution, contrast, and signal-to-noise ratio, and diminishes beam-hardening artefacts.

The disadvantage of CBCT is the limited image quality due to noise and contrast resolution caused by the detection of a large amount of scattered radiation (Figure 1). Of note is the poorer contrast between air and tissue in the CBCT. This is due to a large volume being irradiated during each projected image acquisition and, hence, a large fraction of the beam intensity is attenuated. Some of the scattered radiation falls on the detector, creating background noise that degrades contrast. Hence different areas in the image space may possess different
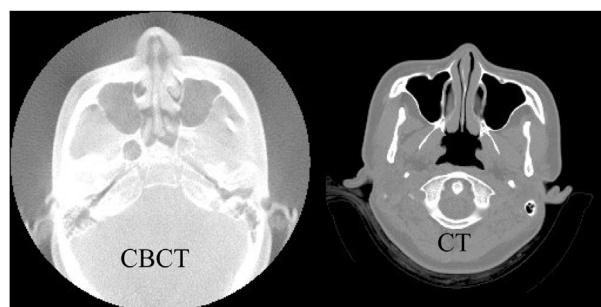

Figure 1. Typical CBCT vs. CT image.

image intensities despite having identical densities in the object space, resulting in image in homogeneity. This complicates the segmentation process, limiting the value of simple techniques like global thresholding and requiring more sophisticated methods.

In conventional CT reconstructions, the voxel grey scale intensity is measured in Hounsfield Units (HUs or CT number) which are uniquely related to the density of the tissue. In CBCT reconstructions, this relationship is not unique: the grey-scale value of each voxel depends on its location in the imaged volume, and thus voxels of similar tissue density but in different locations may have different image values [34] [35] [36]. Accurate absolute density measurement is not possible. There is no standardized system for grey scale value scaling in different CBCT systems. These are arbitrary assigned which complicates scan-to-scan comparison and the universality of the image segmentation approach [37] [38].

1.2. Upper Airway Segmentation from CBCT Images

Upper airway segmentation from CBCT images has received little attention in the literature with the majority of studies limited to sub-sections of the airway and using scans wit contrast and resolution comparable to conventional CT. The purpose of the current work is to identify segmentation techniques suitable for extracting the whole airway from dental-quality CBCT data.

Global thresholding is the simplest segmentation technique. All regions below a chosen pixel intensity (threshold) are identified as air, and all above as tissue. The key decision to be made by the operator or algorithm is the value of the threshold intensity. The Otsu algorithm [39] or variants are often used. Global thresholding performs satisfactorily for many segmentation tasks in conventional CT and MRI images [40]. However the nasal cavity, having a number of intricate and narrow passages and adjacent paranasal sinuses, is susceptible to either under-segmentation (real boundaries between tissue and air are ignored by the segmentation algorithm, and distinct air-filled cavities are merged into one) or over-segmentation (the algorithm returns spurious boundaries where anatomically none exist). Manual segmentation editing, when a trained operator reviews the segmentation result slice-by-slice selecting or deselecting pixels mis-assigned to either air or tissue, is often needed to accurately capture the nasal cavity geometry.

To reduce operator workload, Shi *et al.* [41] [42] proposed an automatic segmentation method for upper airway CBCT images, based on global image thresholding. The threshold value was defined as the one that represented the highest intensity of the region of interest. The segmentation area was limited to the oropharyngeal airway region and maxillary sinuses. The segmentation result was comparable in quality to the ground truth manual segmentation (performed slice-by-slice by a trained operator). The CBCT images used, however, were of high spatial resolution and resolution, comparable to conventional CT images.

Ogawa [43] investigated the use of CBCT for oropharyngeal airway examination in relation to Obstructive Sleep Apnea (OSA). Simple global thresholding

within the Amira segmentation package (Mercury Computer Systems/3D Viz group) was used for airway segmentation for 10 patients with OSA and 10 healthy reference patients. Details of the segmentation approach and segmentation quality were not specified with the main focus being on quantifying airway obstruction.

Farrell [44] reported the automatic segmentation of the oropharynx and laryngopharynx from the soft palate to the vocal cords. The global thresholding technique was applied to segment the airway with the threshold value determined empirically. The selection process was not specified.

Stratemann *et al.* [45] used a combination of global thresholding and manual slice-by-slice editing to segment the airways of 30 subjects for airway shape analysis. This study has the most complete airway coverage (from the nose tip to vocal folds) to date.

Grauer [46] used a user-initialized 3-D surface evolution method [47] for semi-automatic airway segmentation to measure the pharyngeal airway volume and to characterise the shape of the airway. A volume within the airway is selected by the user, and this is grown automatically to fill the airway volume using the image data to control the growth in one of several ways. The method was implemented in Insight SNAP 1.4.0 (Cognitica, Philadelphia, PA & Neuro Image Analysis Laboratories, University of North Carolina, NC). This method evolved into the snake evolution method implemented in ITK-SNAP [47] and used in the present work. The accuracy of the segmentation was evaluated using the coefficient of variation of the volume of the segmented geometries (the ratio of the standard deviation to mean volume), for a subsample of 5 segmentations. The coefficient of variation was less than that for manual segmentation (considered more accurate), but lower than the natural patient-to-patient variability in volume, and thus reliable for the intended purposes.

Bui *et al.* [48] proposed an automated upper airway segmentation method consisting of the Gaussian mixture model (GMM) thresholding and morphological operators to delineate head shape and the active contours used to estimate the region of interest and to initialize the active contour. The initial contour is then evolved using a multi-step level-set segmentation scheme. The results were evaluated using the Dice coefficient [49] [50], the volumetric overlap error (VOE), precision, recall, and F-score measures and compared to the expert manual segmentation.

Salerno *et al.* [51] used a dynamic region-growing algorithm for oropharyngeal airway segmentation. A small volume within the structure of interest is selected. This volume is then grown to encompass pixels that have an intensity similar to that of the original selection. The Otsu algorithm was used to select the threshold value to stop growth. The algorithm was successful, but resulted in incomplete airway models in narrow airways.

Alsufyani *et al.* [52] systematically reviewed the three-dimensional segmentation of the upper airway using CBCT, and noted the lack of a reliable and opti-

mized CBCT protocol for airway imaging and a limited number of published studies into the validity and reliability of 3D airway models generated from CBCT images. Volumetric methods of evaluating the accuracy of segmentation were tested, using manual segmentation and a simplified airway phantom as a reference. The complexity of the nasal cavity and maxillary sinuses, was noted to complicate the segmentation of these regions, compared to the simpler oropharyngeal airway region. Consequently, assessments of the accuracy of segmentation based on studies of the oropharyngeal airway alone are not directly applicable to whole-airway segmentations.

The present work focuses on the segmentation of human airways from CBCT datasets obtained from dental institutions in New Zealand and in the USA. The aim is an objective comparison of the performance of a number of segmentation methods to provide the community with recommendations on how to approach upper airway segmentation. The upper airway, from the tip of the nose to trachea, was split into two parts and segmented from the CBCT images using global thresholding, multi-step level-set, region competition with thresholding, clustering and classification initialisation and edge attraction techniques. The accuracy of the segmentation approaches was assessed using a number of evaluation metrics, such as the Dice Similarity Coefficient, Jaccard Index, Void Fraction, Volumetric Similarity and Volume Ratio, and by direct comparison with manual segmentation.

2. Methodology

An anonymised head and neck cone-beam CT (CBCT) DICOM image series was used (see Figure 2) from a set of 30 studies obtained for the investigation of the upper airway shape variability in adults by Stratemann *et al.* [45]. Ethical approval was granted by the University of Otago Human Ethics Committee. The data was acquired with a New Tom QR DVT 9000 (Aperio, Sarasota, Fla) imaging system. The gender and age of the patients was unknown. Table 1 contains the image characteristics and scan settings used to acquire used CBCT and typical reference CT images used for comparison (acquired using a Siemens Definition CT).

The histogram of the CBCT image series used is plotted in Figure 3 with a histogram of atypical conventional CT image series. Note that the grey scale maxima

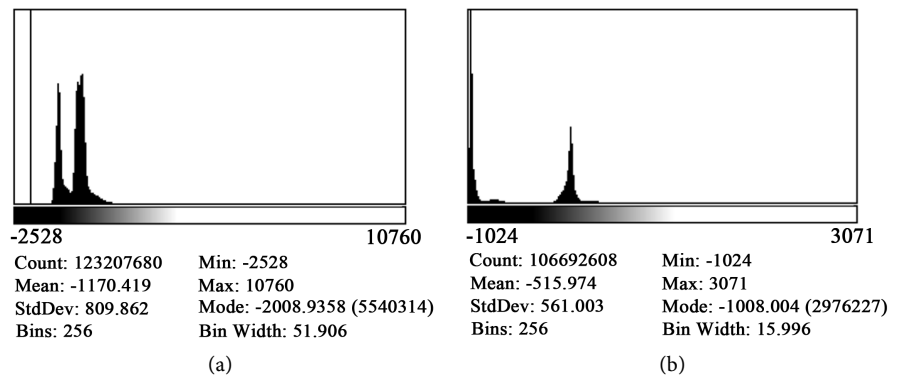**Figure 2.** Sample axial, sagittal and coronal CBCT images.

-2528			10760
Count: 123207680	Min: -2528		
Mean: -1170.419	Max: 10760		
StdDev: 809.862	Mode: -2008.9358 (5540314)		
Bins: 256	Bin Width: 51.906		

(a)

-1024			3071
Count: 106692608	Min: -1024		
Mean: -515.974	Max: 3071		
StdDev: 561.003	Mode: -1008.004 (2976227)		
Bins: 256	Bin Width: 15.996		

(b)

Figure 3. Typical (a) CBCT and (b) conventional CT image intensity histograms.

Table 1. CBCT and reference CT scan setting and image characteristics.

Modality	CBCT	CT
Image resolution	512 × 512 pixels	512 × 512 pixels
Image size	16 bit	16 bit
Slice thickness	0.377 mm	0.6 mm
Pixel spacing	0.377/0.377 mm	0.42/0.42 mm
Reconstruction diameter	193.024 mm	215 mm
Peak Voltage (KVP)	120 kV	140 kV
X-Ray tube current	15 mA	178 mA
Gantry/Detector tilt	0	0
Exposure time	10,000 ms	1000 ms
Patient position	Head First-Supine (HFS)	Head First-Supine (HFS)
Patient orientation	Left-Posterior (L/P)	Left-Posterior (L/P)

and minima and the intensity peaks have different values. This is consistent with the fact that the grey scale values in CBCT do not represent true Hounsfield units. The CT grey scale is in Hounsfield units (HU) which are proportional to tissue density with zero being the radio density of distilled water at standard pressure and temperature (STP) and air having −1000 HU. The CT histogram in Figure 3 shows two distinct peaks, one for air and the other for tissue.

To discriminate air and tissue in CBCT requires the relative position and height of the intensity peaks, rather than absolute number values. The CBCT histogram contains a third peak at low intensity (the very narrow peak at −2008 grey scale units) which corresponds to the region in the image which is excluded from reconstruction, and is consequently empty. Compared to CT, the CBCT image intensity peaks are broader and overlap each other. The air peak is closer to the tissue peak than in CT. These differences are due to the greater noise and the variation of the tissue density-grey scale intensity relationship from region to region in CBCT images, and complicate segmentation.

Scans were split into two regions: nasal cavity (from nares to nasopharynx)

and pharyngeal region (from nasopharynx to larynx) which were segmented separately, and the segmented structures then merged.

2.1. Airway Segmentation

2.1.1. Ground Truth (Manual) Segmentation

CBCT images were manually segmented slice by slice by an experienced user to serve as reference or "ground truth" segmentation (it will be referred to as such, although this is not a true ground truth as the actual airway geometry is not available from any independent source). The segmented 3D model of the airway was stored in STL format. The frontal, maxillary, ethmoidal and sphenoidal sinuses were removed from the airway geometry by manual editing in Meshmixer (Autodesk Inc.) and MeshLab [53]. To preserve finite surface topology, no smoothing was applied.

2.1.2. Semi-Automated Approaches

Three semi-automated approaches were then employed to segment the upper airway from the nares to trachea from the CBCT image: 1) global thresholding, 2) multi-step automated level set segmentation and 3) region competition and edge attraction snakes.

2.1.3. Global Thresholding

A global thresholding procedure was implemented using 3D Slicer, an open source software platform for medical image analysis [54]. The optimum threshold values were estimated by examining the scan histogram. First, the minimum image intensity, I_{max}, corresponding to the image feature of interest was located. Second, the width Δ of the histogram peak at $0.5I_{max}$ was measured. The optimum threshold value was then estimated to be located at around $I_{max} + 0.5\Delta$.

The optimal upper air threshold values were within −550 and −450 grey levels for the images analysed. These varied between image areas due to poor contrast and/or compromised resolution, particularly in the nasal cavity. The nasal passages are narrow (often comparable to pixel size) and intricate and the adjacent paranasal sinuses are particularly susceptible to over- or under-segmentation, when the segmentation result is too fine or too coarse respectively, when using global thresholding. This often results in nasal passages and sinuses being merged and requires laborious and tedious manual slice-by-slice editing to obtain anatomically accurate airway segmentation [27]. Under-segmentation is preferred as it simplifies manual editing in post-processing. Editing the coarse result by adding under-segmented features to the segmentation in 3D Slicer software is more straightforward than removing the spurious boundaries that result from over-segmentation, as the segmentation mask the user is editing does not obscure the raw anatomical data used to guide the operator in tissue differentiation.

A medium threshold value of −500 grey levels was applied to the CBCT images using the Threshold Effect tool within the 3D Slicer Editor module. This value provided best segmentation coverage of the ROI (region of interest), as

determined from visual examination. Segmented images were then used to create a 3D surface model of the airway using the built-in marching cubes-based Model Maker module [55].

2.1.4. Multi-Step Level-Set Segmentation

A multi-step level-set segmentation scheme based on the method of Bui *et al.* [48] was evaluated. The method is based on combining a global and localized region-based active contour. In the first step, a Gaussian mixture model (GMM) is fitted to bone, tissue, and air pixel intensities. The resulting segmentation is then cleaned up using morphological hole-filling. This is used as initialization for a global active contour, in which the active contour evolution equation is given by a Kullback-Leiber (KL) divergence penalty between interior and exterior pixels. Finally, a fine-scale active contour is used, with the energy term favouring the long, thin regions that are present in the upper airway. The advantage of this method is that the global active contour is less sensitive to initialization while the localized active contour is more tolerant of in homogeneity. In addition, anisotropy attempts to overcome the tendency of active contours to shrink surface area of segmented regions. A disadvantage of this method is that detection of bone tissue is poor in the final active contour steps. In addition, the performance of the method is not significantly higher than the classical Chan-Vese level set.

2.1.5. Region Competition and Edge Attraction Segmentation

Region competition snake and edge-based snake evolution segmentation approaches were applied using ITK-SNAP [47], which is an open-source semi-automatic 3D medical image segmentation package.

The initial shape of a 3D surface that represents a segmentation, called a "snake", is coarsely initialized using a number of circular bubbles (the selection of pixels belonging to the region of interest (ROI) being made manually with a cursor) (see for example the pink regions in **Figure 5(i)**). The bubbles may merge as the snake evolves to provide the final segmentation. In region competition, the initialisation is not restricted to being inside the ROI, as the snakes evolve both inward and outward. In contrast, in the edge-based approach the initialization has to be fully contained in the anatomical structure of interest as the snake evolves outward only.

The propagation velocity (distance moved per iteration) of the snake is driven by the snake shape (smoothness) and the by a 'feature' image which is built from the original scan images. The construction of the feature image is a crucial step, because it drives snake evolution and thus may drastically affect the end result. Several types of feature image are possible.

An edge-based feature image is constructed based on the intensity gradient of the original image (computed as the difference in neighbouring pixel intensities). Edges are thus defined as discontinuities in image intensity. The snake is forced to slow down near edges. The edge-based feature image is generated by applying Gaussian blur and computing the magnitude of the image intensity

gradient. The latter is then remapped in decreasing order to the range 0 to 1 to drive the snake evolution.

A region competition feature image forces the snake to attract to boundaries of regions of uniform intensities. These may be produced using thresholding, clustering and classification image intensities to classify image regions as either background or foreground. The feature image is constructed by applying a smoothed step function to edges, with values of around 1 to foreground pixels, values of around −1 to background pixels, and values close to zero to pixels on the border between foreground and background. The feature image resembles a step function (from 1 to −1) which is smoothed to a degree set by the user. The image-dependent velocities force the snake to fill the boundary of the foreground region and to shrink in the background region as the snake iteratively evolves.

2.1.6. Feature Image Construction by Thresholding
An upper threshold value of −468 grey levels was used in a combination with the maximum steepness of the feature function. These were selected based on visual examination of the pre-segmentation result, particularly, the nasal cavity and adjacent sinuses. The aim was to segment as much of the nasal passages while avoiding over-segmentation. The latter may lead to the sinus cavities and nasal cavity structure being artificially merged in the segmented image. The resolution of these anatomical features is often compromised by the narrowness of the air passages in combination with high noise levels in CBCT images. The signal-to-noise ratio (SNR) for the CBCT system used in this study was 26 dB (measured in constancy tests with a calibration phantom).

2.1.7. Feature Image Construction by Clustering or Multi-Thresholding
Multi-thresholding was applied by considering three clusters, each of which is a set of pixels with similar intensities. These clusters identify the parts of the image with intensities corresponding to air, tissue and bone (Figure 5(g) and Figure 5(h)). In this study, 60,000 samples produced the best result, as visually judged by the operator, based on the balance between segmentation completeness and undesirable over-segmentation when different sample sizes were used. No iterations were necessary.

2.1.8. Classification
Three classes of pixels corresponding to air, soft tissue and bone in the image were user-defined within the classification mode of the pre-segmentation to be used for automatic feature image value assignment. 20 - 25 slices from the three axes were used to draw examples of the classes. Particular attention was paid to accurately classifying pixels within the nasal cavity, specifically defining the boundaries between the nasal passages.

2.1.9. Edge Attraction
Edges in the CBCT image were found and accentuated using the edge attraction

mode of image pre-segmentation and snake evolution. The following feature image construction parameters were applied: scale of Gaussian blurring of 1, edge transformation constant of 0.01 and edge transformation exponent of 4 points. These settings provided optimal detection and strengthening of anatomical features boundaries in the images as judged by an experienced user. The snake was initialised within the region of interest, at 20 to 25 locations within the airway passages.

2.2. Metrics Used to Evaluate the Segmentation

The airway geometries extracted by the automatic and semi-automatic segmentation techniques were compared to the "ground truth" manual segmentation using distance-, overlap- and volume-based measures to evaluate the similarity between the ground truth segmentation (GT) and a test segmentation (S), and the completeness of the test segmentation (S).

2.2.1. Distance Metrics

The Hausdorff Distance d_{HD} is defined as the distance from a point in one surface mesh (*i.e.* ground-truth manual segmentation GT) to the closest point in the other mesh (*i.e.* test segmentation S) or vice versa [56]. The maximum of the two one-sided Hausdorff Distances $\text{HD} = \max\{d_{HD}(\text{GT},\text{S}), d_{HD}(\text{S},\text{GT})\}$ was used in this study. HD measures similarity but not completeness. Small values indicate a good match.

2.2.2. Overlap Metrics

The Dice Similarity Coefficient (DSC), Jaccard Index (JAC) and Void Fraction (VF) were used to evaluate both similarity and completeness.

DSC represents the ratio of (twice the intersection (spatial overlap) of the two segmented geometries) to (the sum of the volumes), *i.e.* $\text{DSC} = (2\text{GT} \cap \text{S})/(\text{GT} + \text{S})$ [49] [50]. DSC ranges from 0 to 1, with higher values indicating a higher degree of similarity and completeness between the two geometries. Lower values suggest the test segmentation S is incomplete and/or contains voids not present in GT.

The Jaccard Index represents the intersection between the ground truth and tested segmentation divided by their union [57]. It is related to the DSC by $\text{JAC} = \text{DSC}/(2 - \text{DSC})$. High values are desired, with perfect segmentation returning a value of 1.

The Void Fraction is one minus the intersection of the two geometries divided by the volume of the ground truth geometry, *i.e.* $\text{VF} = (1 - (\text{GT} \cap \text{S})/\text{GT}) \cdot 100\%$. The desired value is 0% indicating a complete overlap between the compared segmentations and absence of voids in the segmented geometry.

2.2.3. Volumetric Measures

The Volume Ratio (VR) and Volumetric Similarity (VS) indicate the degree of possible over-segmentation. They serve a different function to the overlap based metrics which indicate similarity and indirectly, as a reciprocal, the degree of

under-segmentation.

Volume ratio (VR) is the ratio of the segmented volume to the volume of the ground truth segmentation, $VR = S/GT$. The desired value is 1.

The Volumetric Similarity (VS) was defined as the absolute segmentation volume difference divided by the sum of the compared volumes, $VS = (S - GT)/(GT + S)$. The desired value is zero. VS does not measure overlap, and it is possible to obtain a value of zero with two geometries which are different shapes and which do not overlap in any way. As such it should not be used alone.

3. Results

The manual segmentation used as ground truth is shown in **Figure 4(a)**. **Figure 4(b)** shows the same airway after the paranasal sinuses have been manually deleted.

3.1. Pharyngeal Region Segmentation

Figure 5 shows the Hausdorff distance (HD) for the clustering segmentation (the worst performing of the methods which succeeded). The greatest deviations from the ground truth in the lower part of the airway were in three areas: around the pharyngeal recess (Rosenmuller's fossa), around the opening to the Eustachian tube, and around the epiglottis. These locations have voxels within the tissues adjacent to the airway with intensities close to that of with air (such as mucus). Noise in the image also contributes.
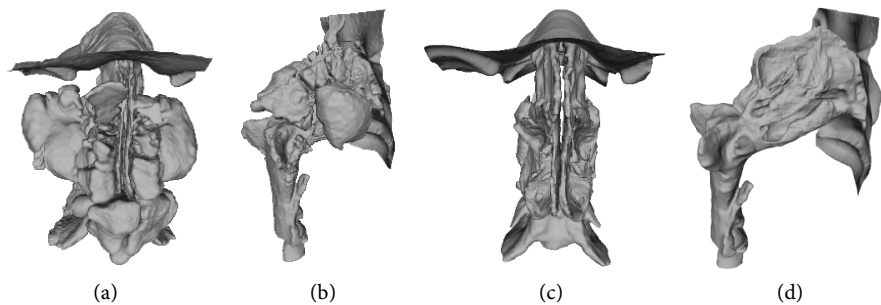

| (a) | (b) | (c) | (d) |

Figure 4. 3D model of the manual ground truth airway segmentation from the CBCT images: (a) before sinus removal; (b) after deletion of sinuses.
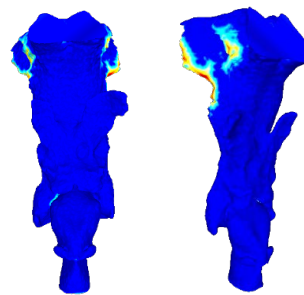

Figure 5. Typical Hausdorff Distance colour map (anterior and lateral view) for the pharyngeal airway segmented by clustering (low and high values are in blue and red respectively).

The root mean square (RMS) Hausdorff distance (HD) values were within 0.4 mm, which is comparable to the voxel size. The maximum HD values for the segmentations using the global thresholding, threshold and classification snake initialisation and multi-step level-set methods ranged from 2.2 to 5.8 mm, with mean values within 0.2 mm (Table 2).

However, for clustering, the HD was at the high end of this range (up to 5 mm, *i.e.* 2.6 times the voxel size) with an RMS of 0.8 mm. This was due to noise in the image, visible in the histogram, resulting in larger clusters.
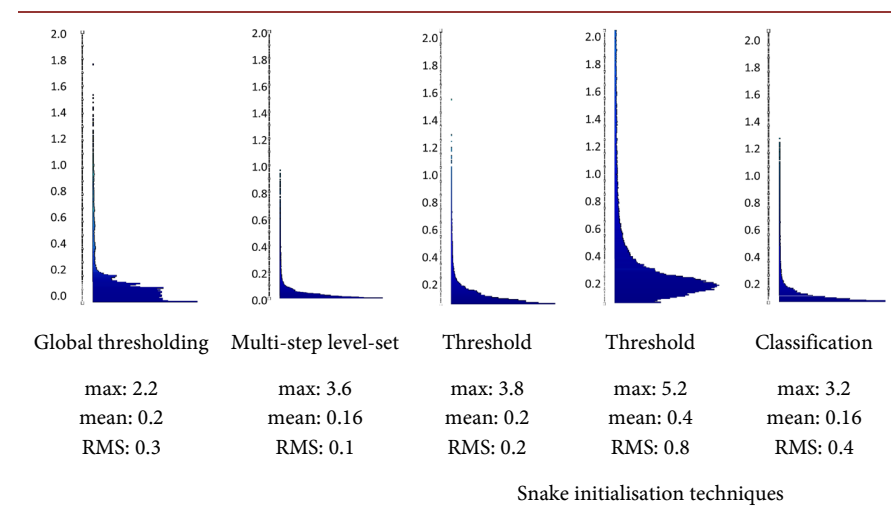
3.2. Nasal Cavity Segmentation

The nasal cavity region was much less straightforward to segment, compared to the pharyngeal region, and was analysed in greater detail.

Mid-axial images (*i.e.* images from the centre of the reconstructed volume), segmented with the global thresholding technique, are shown in Figure 6(a) and Figure 6(b). Examples of the pre-segmentation and final active contour segmentation using the multi-step level-set method are shown in Figure 6(c) and Figure 6(d). The thresholding (used for remapping and feature image construction) and final region competition segmentation for the CBCT data can be seen in Figure 6(e) and Figure 6(f). Figures 6(g)-(j) illustrate clustering, classification and the corresponding final segmentation results.

The edge attraction-based segmentation failed irrespective of the set parameters of active contour initiation and/or segmentation (see Figure 6(k) and Figure 6(l)). The snake leaked into the structures surrounding the airway, eventually encompassing complete ROI.

Figure 7 shows the Hausdorff distance (HD) for the clustering segmentation (the worst performing of the methods which succeeded). The nasal cavity was prone to over- or under-segmentation. The areas problematic for segmentation

Table 2. Hausdorff Distance histograms for the nasopharyngeal-laryngopharyngeal region, for all segmentation methods compared to the ground truth manual segmentation. Vertical scale in millimetres.
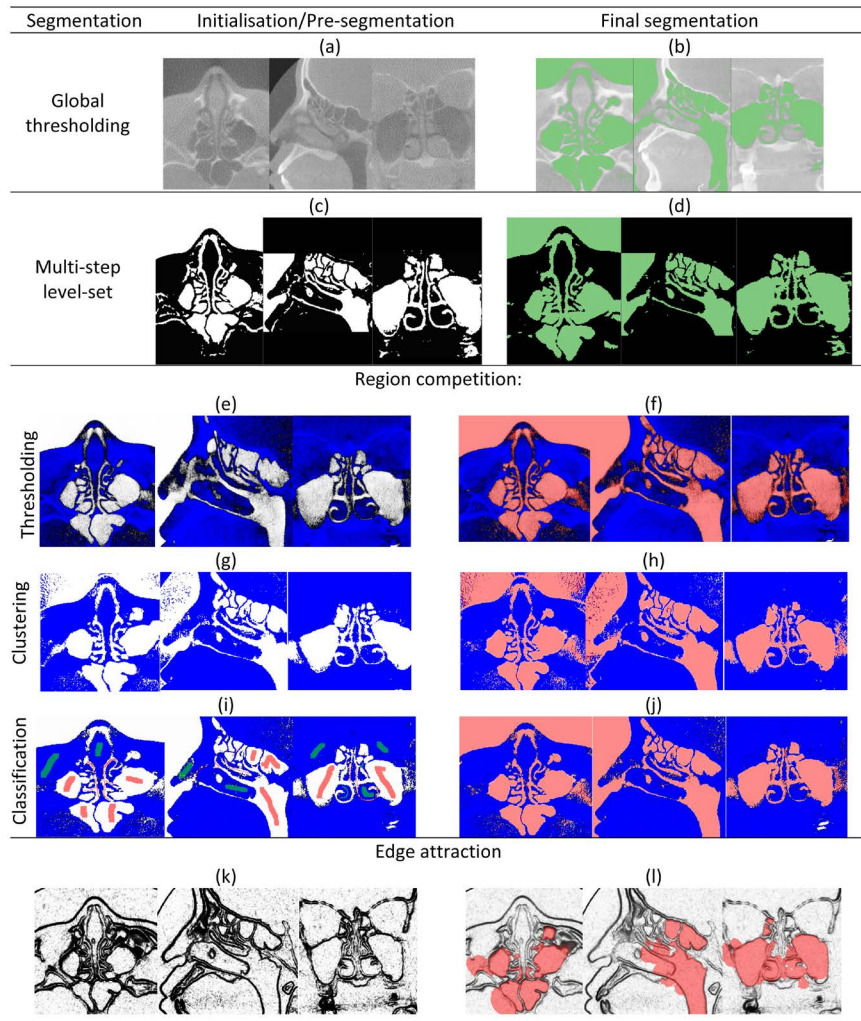

Global thresholding	Multi-step level-set	Threshold	Threshold	Classification
max: 2.2	max: 3.6	max: 3.8	max: 5.2	max: 3.2
mean: 0.2	mean: 0.16	mean: 0.2	mean: 0.4	mean: 0.16
RMS: 0.3	RMS: 0.1	RMS: 0.2	RMS: 0.8	RMS: 0.4

Snake initialisation techniques

Figure 6. CBCT image segmentation of nasal cavity and paranasal sinuses: (a) a set of raw mid-axial, coronal and sagittal CBCT images; (b) Global thresholding label map (segmented airway in green); (c) GMM based segmentation initialisation (in white) and (d) multi-step level-set segmentation result (final segmented airway in pink); (e) Thresholding region competition snake initialisation and (f) resulting final segmentation after region competition; (g) Clustering pre-segmentation and (h) resulting final segmentation after region competition (segmented airway in pink); (i) Classification initialisation with classifiers for air (pink), tissue (green) and bone (blue) and (j) resulting final segmentation after region competition; (k) Edge-based feature image and (l) evolved snake, which clearly failed.
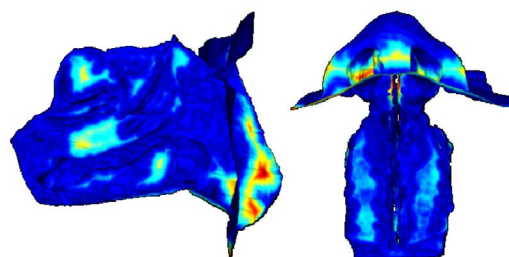

Figure 7. Typical Hausdorff distance colour map (anterior and lateral view) for the nasal cavity segmented by clustering (low values in blue and high in red colour).

were the nasal vestibule and inferior meatus. This was associated with the poor contrast between the air-filled passages and neighbouring tissues of low density such as septal cartilage and mucus.

The maximum Hausdorff distance was 5.0 - 5.6 mm and was similar for all methods (Table 3). Mean and root mean square values ranged from 0.28 to 0.6 mm and 0.47 to 0.91 mm respectively. The highest discrepancies were observed for clustering, around the nasal cavity region in particular. The other techniques all returned a mean HD comparable to the voxel size, which is the minimum that can be expected.

The overlap and volumetric metrics (computed for the nasal cavity only) are shown in Figure 8. The Dice Similarity coefficient ranged from 0.85 for global thresholding to 0.96 for clustering (Figure 8(a)). The other segmentation methods had DSC of 0.90 - 0.92. On this basis alone, clustering would be preferred. The Jaccard Index, being closely related to the DSC, gave similar trends. Global thresholding gave the lowest Jaccard Index of less than 0.75. The highest values, 0.93 and 0.85, were associated with clustering and classification respectively.
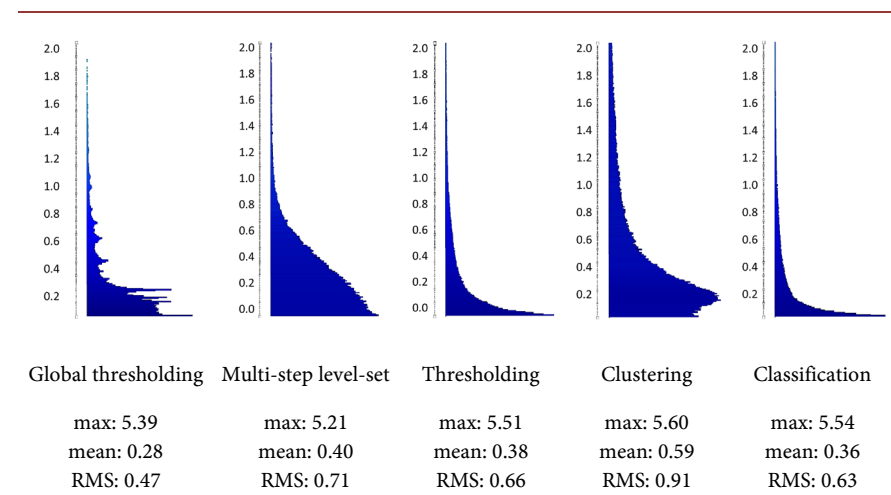
The highest value of the Void Fraction VF (20%) was observed for global thresholding, suggesting this segmentation is incomplete. The lowest VF value of about 5% was observed for clustering, with classification giving the next best of 12%.

The highest values of Volume Ratio (1.5) were observed for clustering (Figure 9(a)). For the other segmentation methods, the nasal cavity volume was within 7% of the ground truth shape.

The Volumetric Similarity (VS) was the lowest for clustering which suggested, along with visual examination, that clustering led to over-segmentation of the nasal cavity (Figure 9(b)).

The value of the key metrics are summarised in Table 4 along with the ideal values. The Jaccard Index is omitted as it conveys no more information than the

Table 3. Hausdorff Distance histograms for the nasal cavity region, for all segmentation methods, compared to the ground truth manual segmentation. Vertical scale in millimetres.
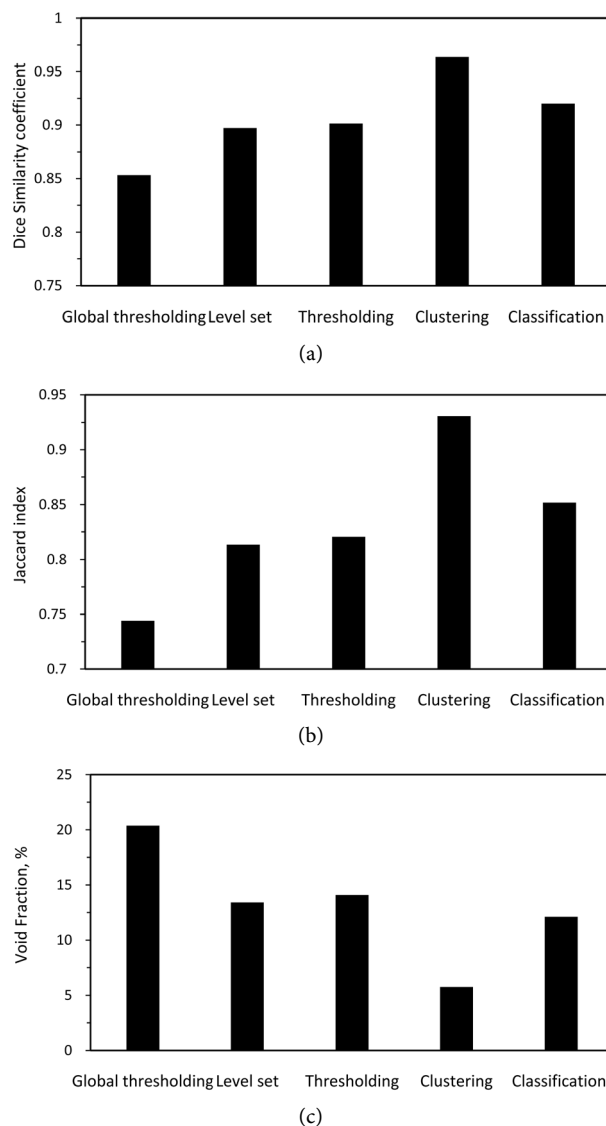

Global thresholding	Multi-step level-set	Thresholding	Clustering	Classification
max: 5.39	max: 5.21	max: 5.51	max: 5.60	max: 5.54
mean: 0.28	mean: 0.40	mean: 0.38	mean: 0.59	mean: 0.36
RMS: 0.47	RMS: 0.71	RMS: 0.66	RMS: 0.91	RMS: 0.63

(a)

(b)

(c)

Figure 8. Dice Similarity Coefficient (a); Jaccard index (b) and Void Fraction (c).

Table 4. Summary of metric values, including the ideal (target) values. The Jaccard Index is omitted as it conveys the same information as the Dice Similarity Coefficient.

|	Mean Hausdorff Distance (nasal cavity) (mm)	Mean Hausdorff Distance (pharyngeal region) (mm)	Dice Similarity Coefficient	Void Fraction (%)	Volume Ratio	Volumetric Similarity
Global Thresholding	0.28	0.2	0.85	20	0.93	1.04
Multi-step level-set	0.40	0.16	0.89	13	0.93	1.04
Thresholding	0.38	0.2	0.90	14	0.98	0.98
Clustering	0.59	0.4	0.96	6	1.53	0.79
Classification	0.36	0.16	0.92	12	1.03	1.01
Ideal value	0	0	1	0	1	0
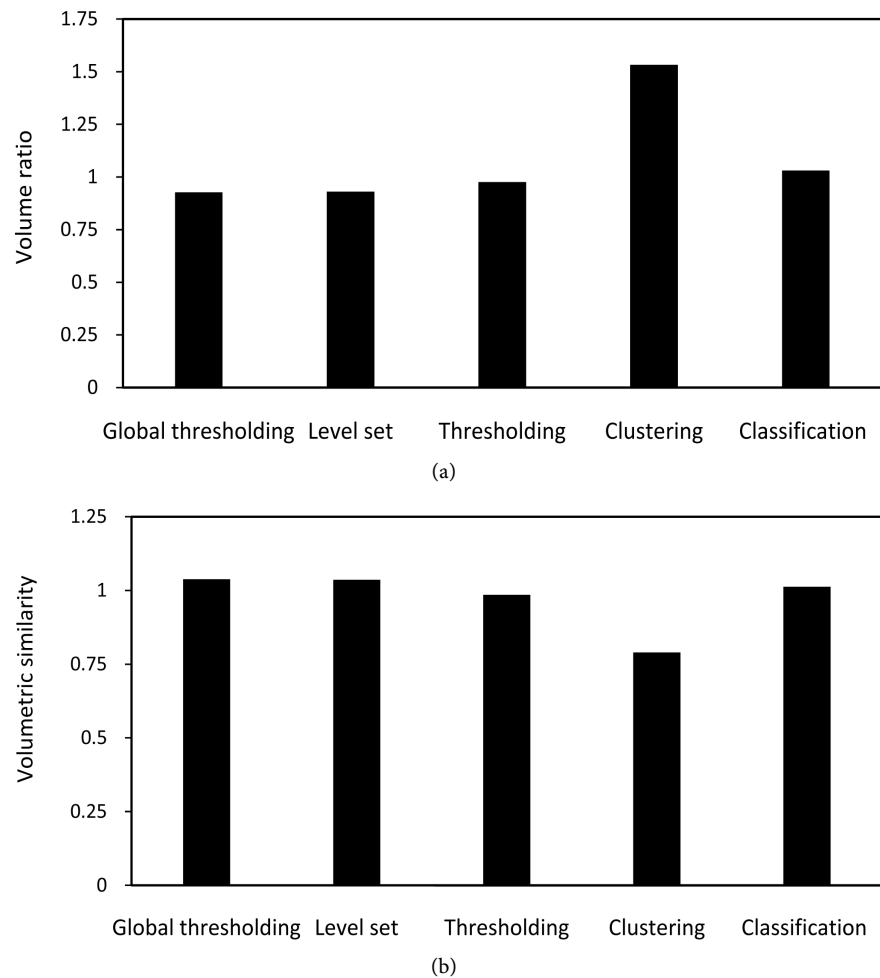
Figure 9. Segmented to ground truth airway volume ratio (a) and volumetric similarity (b).

Dice Similarity Coefficient. These values are computed for the nasal cavity only, and do not include the pharyngeal region, for which discrepancies from the ground truth were much smaller.

4. Discussion

Manual feature segmentation of medical scan data is a laborious and time-consuming task, and there is a demand for an accurate segmentation methods which are either automatic, or require less user input [52]. The accuracy and versatility of the few existing automatic and semi-automatic methods, are under-researched. The global thresholding and snake evolution methods tested here are semi-automatic. Limited manual input was still required for optimum threshold selection, snake initialisation and snake evolution control. The multi-step level-set segmentation algorithm of Bui *et al.* [48] is fully automatic. For all methods, for the typical dental CBCT data segmented here, some manual editing was often needed to fill holes caused by mucus in the nasal cavity being wrongly identified as tissue.

Global thresholding is straight forward but relies heavily on the user's expertise in selecting an optimal threshold value. Use of a single threshold value makes it prone to either under- or over-segmentation. In this study under-segmentation was found to be more favourable when segmenting the nasal cavity due to the complexity of the nasal passages, their narrowness (comparable to pixel size), and the poor image contrast and high noise levels in CBCT images (SNR ~26 dB). The downside of under-segmentation was the presence of voids in the airway geometry. These were time consuming to fix using slice-by-slice manual editing and may compromise the accuracy of the anatomical feature representation when editing is performed on the final 3D mesh. A histogram based approach to optimum threshold estimation, described above, offers a less user-dependent selection method compared to the traditional visual examination method widely employed [43] [44] [45].

The multi-step level-set segmentation algorithm by [48] is fully automatic for selecting regions of interest and segmenting. It did however result in somewhat incomplete segmentation and did not cater for the manual editing mentioned above. However this could be added to a package implementing the method.

The region competition and edge attraction snake evolution methods available in ITK-SNAP give the user a range of semi-automatic tools which reduce user workload. However the quality of the results depended on the user's prior segmentation experience and judgement. Optimal threshold selection remains wholly user-dependent. The sharpness of the feature function affects the degree to which the snake fits the boundary of the foreground region. A steeper feature function was found to reinforce weak edges (those with shallow intensity gradients), of the segmented features: this is of particular use in CBCT images compromised by high noise levels.

The clustering method was found to be unreliable for the CBCT data. It was prone to pick up noise in the image, and led to over-segmentation.

Classification allowed the process to focus on the intensity distribution in the target areas of the image by manually training tissue classifiers. Having more slices used to classify the image into air, tissue and bone improved pixel intensity statistics used for segmentation.

Edge attraction allows for the amount of the blurring applied to be altered. The parameters of the remapping function can also be adjusted to contrast the feature image. Even though edge attraction worked well in the thin homogeneous regions of the image which belong to the same anatomical feature with different intensities, it failed in the nasal cavity region of the CBCT images used. This is due to the low contrast and high levels of noise in CBCT images leading to poorly defined feature edges.

Alsufyani *et al.*, 2012 reported the lack of a protocol to test the accuracy, validity and reputability of the segmented 3D airway models after reviewing 16 studies that applied automatic or semi-automatic techniques for upper airway segmentation from CBCT images. In fact, only 50% of the reviewed studies attempted to validate the segmentation results and reported the methodology

used.

In this study, distance-based, overlap-based and volume-based metrics were used to evaluate CBCT segmentation of the airways. Segmented airway geometries were split into two parts, the nasal cavity and the pharyngeal region. This was done to facilitate the comparison against the ground truth segmentation as the two were recognised to pose different challenges to the segmentation process [45] [48]. The nasal cavity region was more difficult to segment due to the complexity and small scale of the anatomical features. The multi-step level-set segmentation by [48] performed satisfactory in terms of Hausdorff Distance values for the nasopharyngeal-laryngopharyngeal airway region. The method, however, gave higher Hausdorff Distance values for the nasal cavity region (exceeded only by clustering). The measured Dice Similarity Coefficient and Volumetric Similarity, 89% and 1.04 respectively, were lower than those reported by [48], (94% - 97% for DSC and 1.07 - 1.12 for the Volumetric Overlap Error (VOE) which is similar to VS). The latter may be explained by the different quality of the CBCT data used in the present study: Bui *et al.*'s data had finer spatial resolution (0.2 - 0.25 mm in 640 × 640 pixel images).

Region competition snakes with clustering active contour initialisation gave the highest Dice Similarity Coefficient of 97% and the lowest Void Fraction value of 6% suggesting a higher degree of similarity to the ground truth segmentation. This, however, was misleading. The volume ratio was higher (by 50%) than any other method tested, which suggests over-segmentation. Of the range of metrics, clustering returned some results closer to the ideal values than other methods tested, and some farther from the ideal, showing the importance of using more than one metric to evaluate segmentation methods.

Thresholding and classification initialisation methods performed satisfactorily as judged using the spatial distance, overlap and volumetric based evaluation metrics. These, and Multi-step level-set, performed on a par with each other, equally well in all metrics.

Of these, classification performed the best with the highest Dice Similarity Coefficient (92%), Volumetric Similarity closest to 1, the lowest Void Fraction (12%) and Hausdorff Distances bettered only by global thresholding and multi-step level-set. Classification was intuitive to use and did not require extensive segmentation experience. It produced robust results with least amount of manual post-processing required.

The values of the similarity indexes for the nasal cavity airway region were on average lower than those reported in the literature for the nasopharyngeal-laryngopharyngeal airway. Salerno *et al.* [51] reported the mean values of Dice Similarity and Jaccard Coefficients of 94% and 92%, respectively, bettered in this study only by the otherwise erratic clustering method. This suggests that care must to be exercised when segmenting this region of the airway.

An alternative method that uses shape-guided segmentation [58] [59] may have the potential to automate airway segmentation and generate accurate airway from CBCT images, but was not tested here. A representative airway tem-

plate is needed for this method. A population average or mean representative geometry, such as that of [27], may be used as a shape template to match to the airway image being segmented.

5. Conclusions

For segmenting the human upper airway from a typical dental CBCT scan of high noise and low contrast (relative to conventional CT), the following methods were tested:

Global thresholding was the most straightforward, but the optimum threshold value was highly dependent on image quality and the user's familiarity with the anatomy. A simple and objective histogram-based approach performed well for determining the threshold value, compared to subjective visual threshold estimation.

Multi-step level-set segmentation required the least user input, but some input was still required to check the nasal cavity region. Quality indices for the nasal cavity were lower than those reported by Bui *et al.* [48], due to the poorer contrast and resolution of the CBCT images used in the present study.

Snake evolution requires less user input than global thresholding, but more than multi-step level-set. We recommend using the minimum remapping function steepness setting for noisy CBCT data, due to the shallow intensity gradients (weak edges).

Several methods were trialled to initialize the snake, which showed markedly different performance:

Initialization with clustering was compromised by the noise in the images and performed poorly for the nasal cavity region.

Initialization with classification was the most promising method. The more scan slices used to train the air, tissue and bone classifiers, the better the final result.

Initialization by edge attraction failed in the nasal cavity region due to the low contrast and high noise in the CBCT images. Use of blurring may return better performance for images of higher contrast and higher signal-to-noise ratios than used here.

Manual editing was required to fill voids due to mucus for all methods tested here.

Several distance, overlap, and volumetric metrics were trialled to evaluate the quality of the segmentation objectively. The use of any single metric can be misleading, particularly when the data is under- or over-segmented.

The nasal cavity is particularly challenging to segment due to the small scale of the airway passages, the intricate structure, the adjacent paranasal sinuses, limited spatial resolution, low contrast and noise. The greatest discrepancies between the segmentation methods tested and a careful manual segmentation were about 5 mm (Hausdorff distance) and were found in the nasal vestibule and inferior meatus, due to the septal cartilage and mucus often being incorrectly seg-

mented as part of the airway. The pharyngeal region was relatively insensitive to the segmentation method. The pharyngeal recess (Rosenmuller's fossa) was the most difficult part of this region to segment correctly.

Acknowledgements

References

[1] WHO Strategy for Prevention and Control of Chronic Respiratory Diseases, WHO/MNC/CRA/02.1, 2002.

[2] Beachey, W. (2012) Respiratory Care Anatomy and Physiology. 3rd Ed., Foundations for Clinical Practice, Elsevier, Amsterdam, Netherlands.

[3] Churchill, S.E., Shackelford, L.L., Georgi, J.N., Black, M.T. (2004) Morphological Variation and Airflow Dynamics in the Human Nose. *American Journal of Human Biology*, **16**, 625-638. https://doi.org/10.1002/ajhb.20074

[4] Noback, M.L., Harvati, K. and Spoor, F. (2011) Climate-Related Variation of the Human Nasal Cavity. *American Journal of Physical Anthropology*, **145**, 599-614. https://doi.org/10.1002/ajpa.21523

[5] Franciscus, R.G. and Long J.C. (1991) Variation in Human Nasal Height and Breadth. *American Journal of Physical Anthropology*, **85**, 419-427. https://doi.org/10.1002/ajpa.1330850406

[6] Eccles, R. (1996) A Role for the Nasal Cycle in Respiratory Defence. *European Respiratory Journal*, **9**, 371-376. https://doi.org/10.1183/09031936.96.09020371

[7] Kim, S.K. and Chung, S.K. (2004) An Investigation on Airflow in Disordered Nasal Cavity and Its Corrected Models by Tomographic PIV. *Measurement Science and Technology*, **15**, 1090-1096. https://doi.org/10.1088/0957-0233/15/6/007

[8] Hörschler, I., Brücker, C., Schröder, W. and Meinke, M. (2006) Investigation of the Impact of the Geometry on the Nose Flow. *European Journal of Mechanics-B/ Fluids*, **25**, 471-490.

[9] Ishikawa, S., Nakayama, T., Watanabe, M. and Matsuzawa, T. (2006) Visualization of Flow Resistance in Physiological Nasal Respiration: Analysis of Velocity and Vorticities using Numerical Simulation. *Archives of Otolaryngology-Head & Neck Surgery*, **132**, 1203-1209. https://doi.org/10.1001/archotol.132.11.1203

[10] Jeong, S.-J., Kim, W.-S. and Sung, S.-J. (2007) Numerical Investigation on the Flow Characteristics and Aerodynamic Force of the Upper Airway of Patient with Obstructive Sleep Apnea using Computational Fluid Dynamics. *Medical Engineering and Physics*, **29**, 637-651.

[11] Mylavarapu, G., Murugappan, S., Mihaescu, M., Kalra, M., Khosla, S. and Gutmark, E. (2009) Validation of Computational Fluid Dynamics Methodology Used for Human Upper Airway Flow Simulations. *Journal of Biomechanics*, **42**, 1553-1559.

[12] Gambaruto, A., Taylor, D. and Doorly, D. (2009) Modelling Nasal Airflow using a Fourier Descriptor Representation of Geometry. *International Journal for Numerical Methods in Fluids*, **59**, 1259-1283. https://doi.org/10.1002/fld.1866

[13] Taylor, D., Doorly, D. and Schroter, R. (2010) Inflow Boundary Profile Prescription for Numerical Simulation of Nasal Airflow. *Journal of the Royal Society Interface*, **7**,

515-527. https://doi.org/10.1098/rsif.2009.0306

[14] Lee, J.-H., Na, Y., Kim, S.-K. and Chung, S.-K. (2010) Unsteady Flow Characteristics through a Human Nasal Airway. *Respiratory Physiology and Neurobiology*, **172**, 136-146.

[15] Schroeter, J.D., Garcia, G.J.M. and Kimbell, J.S. (2010) A Computational Fluid Dynamics Approach to Assess Interhuman Variability in Hydrogen Sulfide Nasal Dosimetry. *Inhalation Toxicology*, **22**, 277-286.
https://doi.org/10.3109/08958370903278077

[16] Kumar, H., Spence, C.J.T. and Tawhai, M.H. (2015) Modeling the Pharyngeal Pressure during Adult Nasal High Flow Therapy. *Respiratory Physiology and Neurobiology*, **219**, 51-57.

[17] Kumar, H., Jain, R., Douglas, R.G. and Tawhai, M.H. (2016) Airflow in the Human Nasal Passage and Sinuses of Chronic Rhinosinusitis Subjects. *PLoS ONE*, **11**, e0156379.

[18] Van Hove, S., Storey, J., Adams, C., Dey, K., Geoghegan, P., Kabaliuk, N., Oldfield, S., Spence, C., Jermy, M. and Suresh, V. (2016) An Experimental and Numerical Investigation of CO_2 Distribution in the Upper Airways during Nasal High Flow Therapy. *Annals of Biomedical Engineering*, **44**, 3007-3019.

[19] Park, K., Brucker, C. and Limberg, W. (1997) Experimental Study of Velocity Fields in a Model of Human Nasal Cavity by DPIV. *Proceedings 7th International Conference on Laser Anemometry, Advances and Applications*, Karlsruhe, 617-626.

[20] Kelly, J., Prasad, A. and Wexler, A. (2000) Detailed Flow Patterns in the Nasal Cavity. *Journal of Applied Physiology*, **89**, 323-337.

[21] Kook Kim, J., Yoon, J.-H., Hoon, K.C., Wook Nam, T., Bo Shim, D. and Ae Shin, H. (2006) Particle Image Velocimetry Measurements for the Study of Nasal Airflow. *Actaoto-laryngologica*, **126**, 282-287. https://doi.org/10.1080/00016480500361320

[22] Spence, C., Buchmann, N. and Jermy, M. (2012) Unsteady Flow in the Nasal Cavity with High Flow Therapy Measured by Stereoscopic PIV. *Experiments in Fluids*, **52**, 569-579. https://doi.org/10.1007/s00348-011-1044-z

[23] Spence, C., Buchmann, N., Jermy, M. and Moore, S. (2011) Stereoscopic PIV Measurements of Flow in the Nasal Cavity with High Flow Therapy. *Experiments in Fluids*, **50**, 1005-1017. https://doi.org/10.1007/s00348-010-0984-z

[24] Geoghegan, P., Buchmann, N., Spence, C., Moore, S. and Jermy, M. (2012) Fabrication of Rigid and Flexible Refractive-Index-Matched Flow Phantoms for Flow Visualisation and Optical Flow Measurements. Experiments in Fluids, 52.

[25] Liu, Y., Johnson, M.R., Matida, E., Kherani, S. and Marsan, J. (2009) Creation of a Standardized Geometry of the Human Nasal Cavity. *Journal of Applied Physiology*, **106**, 784-795. https://doi.org/10.1152/japplphysiol.90376.2008

[26] Lee, C.F., Abdullah, M.Z., Ahmad, K.A. and LutfiShuaib, I. (2013) Standardization of Malaysian Adult Female Nasal Cavity. Computational and Mathematical Methods in Medicine.

[27] Nejati, A., Kabaliuk, N., Jermy, M. and Cater, J. (2016) A Deformable Template Method for Describing and Averaging the Anatomical Variation of the Human Nasal Cavity. *BMC Medical Imaging*, **16**, 55.
https://doi.org/10.1186/s12880-016-0154-8

[28] Gambaruto, A.M., Taylor, D.J. and Doorly, D.J. (2012) Decomposition and Description of the Nasal Cavity Form. *Annals of Biomedical Engineering*, **40**, 1142-1159. https://doi.org/10.1007/s10439-011-0485-0

[29] Bushberg, J.T., Seibert, J.A., Leidholdt, E.M. and Boone, J.M. (2012) The Essential Physics of Medical Imaging. 3rd Edition, Lippincott Williams and Wilkins, Philadelphia.

[30] Boas, F.E. and Fleischmann, D. (2012) CT Artifacts: Causes and Reduction Techniques. *Imaging in Medicine*, **4**, 229-240.

[31] Loubele, M., Bogaerts, R., Van Dijck, E., Pauwels, R., Vanheusden, S., Suetens, P., Marchal, G., Sanderink, G. and Jacobs, R. (2008) Comparison between Effective Radiation Dose of CBCT and MSCT Scanners for Dentomaxillofacial Applications. *European Journal of Radiology*, **71**, 461-468.

[32] Orth, R.C., Wallace, M.J. and Kuo, M.D. (2008) C-Arm Cone-Beam CT: General Principles and Technical Considerations for Use in Interventional Radiology. *Journal of Vascular and Interventional Radiology*, **19**, 814-820.

[33] Scarfe, W.C., Farman, A.G. and Sukovic, P. (2006) Clinical Applications of Cone-Beam Computed Tomography in Dental Practice. *Journal of the Canadian Dental Association*, **72**, 75-80.

[34] Swennen, G.R. and Schutyser, F. (2006) Three-Dimensional Cephalometry: Spiral Multi-Slice vs. Cone-Beam Computed Tomography. *American Journal of Orthodontics and Dentofacial Orthopedics*, **130**, 410-416.

[35] Armstrong, R.T. (2006) Acceptability of Cone Beam CT vs. Multi-Detector CT for 3D Anatomic Model Construction. *Journal of Oral and Maxillofacial Surgery*, **64**, 37.

[36] De Vos, W., Casselman, J. and Swennen, G.R. (2009) Cone-Beam Computerized Tomography (CBCT) Imaging of the Oral and Maxillofacial Region: A Systematic Review of the Literature. *International Journal of Oral and Maxillofacial Surgery*, **38**, 609-625.

[37] Mah, P., Reeves, T.E. and McDavid, W.D. (2010) Deriving Hounsfield Units using Grey Levels in Cone Beam Computed Tomography. *Dentomaxillofacial Radiology*, **39**, 323-335. https://doi.org/10.1259/dmfr/19603304

[38] Rezi, T., Niknami, M. and Chazani, F.A. (2014) Relationship between Hounsfield Unit in CT Scan and Grey Scale in CBCT. *Journal of Dental Research, Dental Clinics, Dental Prospects*, **8**, 107-110.

[39] Otsu, N. (1979) A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, **9**, 297-302. https://doi.org/10.1109/TSMC.1979.4310076

[40] Bankman, I.N. (2000) The Handbook of Medical Imaging: Processing and Analysis. Academic Press.

[41] Shi, H., Scarfe, W.C. and Farman, A.G. (2006) Maxillary Sinus 3D Segmentation and Reconstruction from Cone Beam CT Data Sets. *The International Journal for Computer Assisted Radiology and Surgery*, **1**, 83-89. https://doi.org/10.1007/s11548-006-0041-9

[42] Shi, H., Scarfe, W.C. and Farman, A.G. (2006) Upper Airway Segmentation and Dimensions Estimation from Cone-Beam CT Image Datasets. *The International Journal for Computer Assisted Radiology and Surgery*, **1**, 177-186. https://doi.org/10.1007/s11548-006-0050-8

[43] Ogawa, T., Enciso, R., Shintaku, W.H. and Clark, G.T. (2007) Evaluation of Cross-Section Airway Configuration of Obstructive Sleep Apnea. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, **103**, 102-108.

[44] Farrell, M.L. (2009) Fully Automated Upper Airway Segmentation and Surfacing on a GPU from Cone-Beam CT Volumes. Thesis, Ohio University.

[45] Stratemann, S., Huang, J.C., Maki, K., Hatcher, D. and Millere, A.J. (2011) Three-Dimensional Analysis of the Airway with Cone-Beam Computed Tomography. *American Journal of Orthodontics and Dentofacial Orthopedics*, **140**, 607-615.

[46] Grauer, D. (2007) Airway Volume and Shape from Cone-Beam CT: Relationship to facial Morphology. University of North Carolina, Thesis.

[47] Yushkevich, P.A., Piven, J., Hazlett, H.C., Smith, R.G., Ho, S., Gee, J.C. and Gerig, G. (2006) User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability. *Neuroimage*, **31**, 1116-1128.

[48] Bui, N.H., Ong, S.H. and Foong, K.W.C. (2015) Automatic Segmentation of the Nasal Cavity and Paranasal Sinuses from Cone-Beam CT Images. *International Journal of Computer Assisted Radiology and Surgery*, **10**, 1269-1277.

[49] Dice, L.R. (1945) Measures of the Amount of Ecologic Association between Species. *Ecology*, **26**, 297-302. https://doi.org/10.2307/1932409

[50] Sørensen, T. (1948) A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species and Its Application to Analyses of the Vegetation on Danish Commons. *Kongelige Danske Videnskabernes Selskab*, **5**, 1-34.

[51] Salerno, S., Gagliardo, C., Vitabile, S., Militello, C., La Tona, G., Giuffrè, M., Lo Casto, A. and Midiri, M. (2014) Semi-Automatic Volumetric Segmentation of the Upper Airways in Patients with Pierre Robin Sequence. *The Neuroradiology Journal*, **27**, 487-494.

[52] Alsufyani, N., Flores-Mir, C. and Major, P. (2012) Three-Dimensional Segmentation of the Upper Airway using Cone Beam CT: A Systematic Review. *Dentomaxillofacial Radiology*, **41**, 276-284. https://doi.org/10.1259/dmfr/79433138

[53] Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F. and Ranzuglia, G. (2008) MeshLab: an Open-Source Mesh Processing Tool. 6*th Eurographics Italian Chapter Conference*, 129-136.

[54] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C. and Pujol, S. (2012) 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*, **30**, 1323-1341.

[55] Lorensen, W.E. and Cline, H.E. (1987) Marching Cubes: A High Resolution 3D Surface Construction Algorithm.

[56] Aspert, N., Santa-Cruz, D. and Ebrahim, T. (2002) Mesh: Measuring Errors between Surfaces using the Hausdorff Distance. In: *Proceedings of the IEEE International Conference on Multimedia and Expo*, Vol. 1, 705-708. https://doi.org/10.1109/ICME.2002.1035879

[57] Levandowsky, M. and Winter, D. (1971) Distance between Sets. *Nature*, **234**, 34-35. https://doi.org/10.1038/234034a0

[58] Borenstein, E. and Malik, J. (2006) Shape Guided Object Segmentation. *Proceedings of the* 2006 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 1969-976. https://doi.org/10.1109/CVPR.2006.276

[59] Wang, H. (2012) Shape-Guided Interactive Image Segmentation. University of Alberta, Thesis.