Dynamic data analysis for pediatric airways

Chun-Wei Liu

December 4, 2014

Abstract

The analysis of pediatric airway geometry using computed tomography (CT) images has provided rich diagnostic cues for doctors. Recently, dynamic CT data has started to provide a better characterization of pediatric airways throughout the breathing cycle, for example to assess tracheomalacia. However, how to preprocess the increasing amount of dynamic data and how to analyze these data are still open questions. In this work, I performed dynamic data analysis using computer vision and machine learning approaches on synthetic and real airway data. In the future, I aim at building a 4D atlas for pediatric airways and at extending the approaches to other dynamic data modalities.

1 Introduction

The analysis of pediatric airway geometry using 3D computed tomography (CT) images has provided rich cues for doctors to diagnose respiratory issues for patients. Both radiologists and physicians have been working on the field for a while. In the image analysis side over the past decade, radiologists started to propose the algorithm for measuring airway lumenal area using multidetector row CT [1]. In the other respects, physicians adopted benefits from imaging analysis for their clinical studies. For example, how their patients with airway issues were recovered after surgeries [2]. The progress made in image analysis these days was making the communication of both parties much easier by providing more informative statics from images. For instance, given a CT image from a subject with some manually annotative landmarks, a machine is able to learn from CT images of normal control data to build a subject-specific control atlas (a mean statics from the population of the subject.) Then machine can provide statics for where the positions have smaller cross-sectional area to doctors which makes diagnosing tracheal stenosis more actuary [3].

What can we learn from a four-dimensional CT (4D) image, an image set contains up to 16 3D CT images with respiratory motion induced image changes across the set that are not available on a single-component 3D CT image? Compared to spirometry, analysis on 3D CT images provides additional information about the likely *location* causing respiratory distress. Furthermore, 4D CT images can provide information about *when* the respiratory issue occurs during the breathing cycle. Four-dimensional CT imaging was developed to provide an estimate of tumor motion for radiotherapy treatment planning [4]. After then, a method for dynamic ventilation imaging of the full respiratory cycle from 4D CT images was developed [5]. Recently, 4D CT imaging has provided a better characterization of pediatric airways throughout the breathing cycle, for example to assess tracheomalacia, which is a disease causing a temporary collapse of the partial airway.

However, how to preprocess the increasing amount of dynamic data and how to analysis these data for pediatric airways are still open questions. While standard techniques to analyze 3D CT data could be applied for the analysis of 4D CT data in a frame-by-frame fashion, the following two major challenges need to be addressed: First, the cost of manual annotations or preprocessing increases proportionally to the number of CT scans in a breathing cycle. For example, a 4D CT data, including a set of 8 to 16 3D CT frames, took about 10 minutes for preprocessing each frame. Manually doing these preprocessing is time consuming. Second, by not sharing information between time steps of the breathing cycle temporal inconsistence may be caused for the analysis. In this work, I am going design an automatic 4D data processing framework to address the above issues by performing dynamic data analysis using computer vision and machine learning approaches.
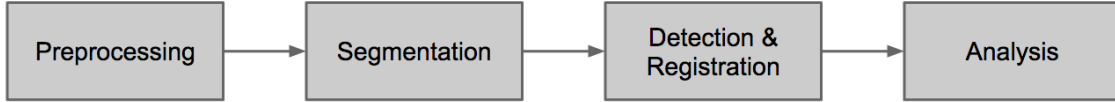
Figure 1: Illustration of the proposed 4D CT image processing framework. From CT image preprocessing, image segmentation, landmark detection and functional data registration, to final analysis.

In Section 2, I will introduce the framework applied to pediatric airway analysis. Experiments on real airway data are shown in Section 3. Section 4 discusses the future works, including building a 4D atlas for pediatric airways and extending the approaches to other dynamic data modalities.

2 Methods

To perform data analysis for subjects with large spatial differences, spatial normalization in the form of registration is an important step. Registration can be approached in different ways. First, image registration, which uses image intensities as features, has been developed in medical image analysis for the past decade. Most of these methods have to do with defining an energy function with data term (dissimilarity between source image and target image) and regularization term (atypicality of mapping from source image to target image), and then develop optimization method to find a best mapping which minimizes such energy. See the review articles by Hill et al. and Sotiras et al. on this topic [6, 7] for details. For particular application for registration on one subject among a short time period, Guerrero et al. developed a deformable algorithm for dynamic ventilation imaging from 4D CT [5].

On the other hand, shape registration, which uses geometric cues as major features, has succeeded in many applications in computer vision and computer graphics [8, 9]. With the same fashion of energy minimization framework as image registration, these methods rely on shape features, e.g. curvature or level of bending, to define the data term. However, using such techniques to align pediatric airway data is still difficult. Hong et al. proposed a simplified airway model which allows for much easier registration for further analysis [3]. Figure 7 shows the anatomy I am looking at in this work. I started this project by extending Hong et al.'s simplified airway model to a 4D CT processing framework. Figure 1 illustrated the automatic 4D CT processing framework.

2.1 Image preprocessing and segmentation

The algorithm segmented the airway from 3D CT image using Otsu's method and two manually chosen seeds that bracket the airway. For better adaptive to current medical image processing libraries, the framework starts with reformatting data from DICOM images to NRRD images. Then a padding and filtering operator for making the image boundary has the same intensity as air is applied on NRRD images. With scripts automatically applying preprocessing programs reduced costs of manually operating medical image softwares, such as Slicer or ParaView.

Otsu's method finds a threshold to separate data into two clusters, so as to try to make each cluster as tight as possible. This is achieved by maximizing the inter-subject variances and minimizing intra-subject variances in terms of voxel intensity. In our case for segmenting airway, the two clusters were air and non-air regions. To remove the air regions that were not in the airway, morphological operator closing (with a sphere structuring pattern) first be applied to remove the actual airway. This computed a map of air regions outside the subject. We can apply the map as a filter to remove the air outside the airway to get a clean airway in the subject. Then, two seeds further help to extract trachea and exclude bronchus in the airway.

With this segmentation, the airway can be approximated by statistical models [10]. Focus on measuring the size of airway, we used a centerline with cross sections to represent the airway. The centerline is inferred based on the heat distribution along the airway flow that is solved for by a Laplace equation. Cross sections
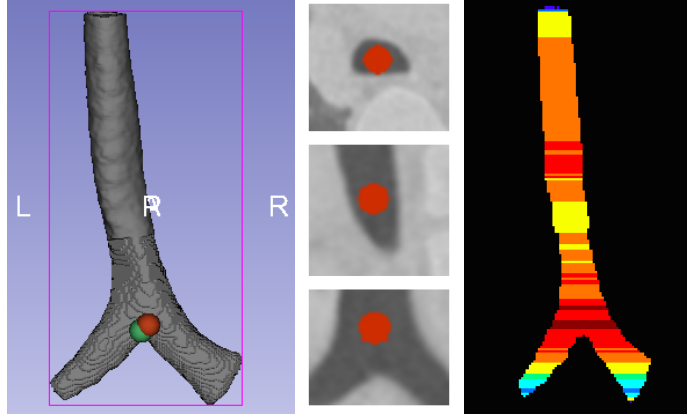
Figure 2: Visualization of the framework of landmark detection. **Left.** First, an airway geometry is segmented by Otsu-thresholding. Green marker is the ground truth annotation and red marker is the predicted location of TC. **Middle.** Second, CHOG features are computed in the center of the trachea in each location along the airway. From top to down: the Axial, sagittal, and coronal plane. **Right.** Finally, visualization of the probabilities of applying the trained classifier on these different hypotheses. Dark red indicates the highest likelihood of TC. In this case, the predicted TC is only 1.97 mm (the total length of the airway is 159 mm) away from the ground truth annotation. This error is only 1.2% of the entire airway.

are cut from segmented airway geometry using planes that are orthogonal to the centerline. The area of the cross sections along the airway is then used as a 1D functional data representation of an airway.

2.2   Landmark detection

We can register functional data using some common landmarks across subjects. Typically landmarks are determined manually. To reduce the manual annotation cost, based on Dalal and Triggs's object detection framework [11], I propose a landmark detection framework using concatenated Histogram of Gaussian (CHOG) features and a geometric prior. Figure 2 illustrates the detection of the trachea carina (TC). The goal of the landmark detection is to capture the dynamics of landmarks. Therefore, the accuracy of landmark detection should be high enough to compensate the dynamics in the 4D CT image. If the prediction error is larger than the actual dynamics in the 4D CT image, then the predicted location cannot be determined.

HOG is well designed normalized local histograms of image gradient orientation in a dense sample grid. The original purpose of this feature was for human detection [11]. Given the popularity of HOG based object detectors, researchers even tried to answer why it works (or did not work) by visualizing the feature spaces [12]. However, it has more general applicability as it captures edge or gradient structures that are very characteristic of local shape. It can also be efficiently computed. In practice the HOG is implemented by dividing an image window into small spatial cells, for each cell accumulating a histogram of gradient directions with different bins over pixels of the cell. The concatenated histogram over cells would be the final representation of an image window. To apply HOG on 3D images, instead of computing histograms in arbitrary 3D orientations, I computed 2D HOG in the axial, coronal, and sagittal planes, which are the three perspectives for user annotations, and concatenated them together as a three times longer concatenated histogram (CHOG). In this work, I used this simple trick to generalize HOG from 2D to 3D.

The first step of the landmark detection framework then is to train a binary classifier using CHOG. For doing high dimensional data and low sample size statical analysis, Distance Weighted Discrimination can be applied to improve the training [13]. Yet, for simplistic and speed, linear Support Vector Machine (SVM) is used as a baseline throughout study. In the prediction stage, the trained classifier can then be applied to the detect the particular landmark. Higher score generally implies a higher likelihood of a hypothesis.
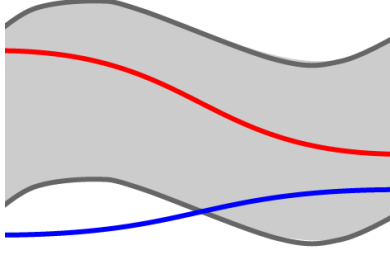
3

Figure 3: Visualization of the band depth. The red curve fully belongs to the shaded region, but the blue curve is partially belonged to the shaded region.

After landmarks are located, I registered the functional data using b-spline curve registration [14]. The technique needs at least four landmarks for registration. In Hong et al.'s original paper, each subject had five visible landmarks: from nasal spine, choana, epiglottis tip, true vocal cord (TVC), and TC. In our dynamic data, most of subjects only have TVC and TC due to a limited of view. So we can only align these curves by fixing the start and the end of the curve and apply linear interpolation. In some cases, TC is even the only available landmark which makes alignment impossible using linear interpolation. I therefore apply a heuristic assumption to these special cases. Given we have decent amount of data, the assumption is that a given subject has the same length of trachea from TVC to TC as the most relevant subject (in terms of age) in our dataset. Therefore, we can compute the proportion of the imaged trachea by measuring the ratio of the length of the current trachea in physical space and the length of the most relevant subject from TVC to TC.

2.3 Statistical atlas analysis

Given spatially aligned functional data I would like to capture population changes with respect to some factors, say age. We can use a blending kernel applying on the dataset to interpret a subject with the target age. For example, I used a Gaussian weight function $w_i(a_i; \sigma, \bar{a}) = c \exp{(a_i - \bar{a})/2\sigma^2}$, where a_i is the age for the observation i, σ is a predefined standard deviation and c is the normalization constant for fitting data to a specific age.

To further analyze the weighted data, I applied weighted functional boxplots to build a statistical atlas for each dynamic subject [15]. Functional boxplots were originally proposed by Sun and Genton [16] based on the definition of band depth for functional data.

Band depth is used to rank functional data for ordering it from the center outward. To order functional data, we first define a subset and the membership of a data in the subset. Given a functional dataset, any two function curves in the dataset can form a set in which the region between these two curves belong. A curve fully belongs in this subset if it is in the region from the beginning to the end of the function. A curve might be partially belong in the subset. Figure 3 shows the difference. A band depth of a functional data curve can be defined as the ratio of the number of subset in which data fully belongs over all possible subsets.

Given a set of functional data $Y = \{y_i | i = 1, ..., n\}$, a combinatorial function C which enumerates all two pair combinations in a set, and a band function $B(y_1, y_2) = \{(t, x(t)) : t \in \mathrm{T}, \min(y_1(t), y_2(t)) \leq x(t) \leq \max(y_1(t), y_2(t))\}$, the band depth D of a functional data y with respect to a set Y can be defined as

$$D(y; Y) = \sum_{y_i, y_j \in C(Y)} I[y \subset B(y_i, y_j)], \tag{1}$$

where I is an indicator function. A generalized version of band depth is the weighted modified band depth

4

$$D'(y; Y) = \sum_{y_i, y_j \in C(Y)} w_i w_j \lambda[B(y_i, y_j)] \qquad (2)$$

where λ is the Lebesgue measure, and ws are the weights of data for fitting the population. If a curve y is only partially in a subset $B(y_i, y_j)$, the λ measures the ratio of time that y is in the subset $B(y_i, y_j)$. If a curve is fully in the subset, the λ measurement is one. I applied (2) to compute a population atlas and to plot subject dynamics upon the population atlas using (1).

When a rank of functional data is available, we can compute interesting statistics such as median, interquartile range, and outliers of the population. For example, given a set of dynamic data S and a normal control atlas N, to identify an abnormal airway collapse, we can apply point estimator on the functional data that is out of the normal control atlas. The maximum "Collapsed dynamics" can be defined as

$$E(S, N) = \frac{\max_x(s_x) - \min_x(s_x)}{\max_x(s_x)}, \qquad (3)$$

where $x = \{i, s_i < n_i | s \in S, n = \inf N\}$

3 Experiments

I applied the proposed framework on real world dynamic data. This section will provide implementation details and parameter settings, and will also provide results and observations from experiments.

3.1 Landmark detection results

First I conducted a preliminary experiment before further developing landmark detection techniques by answering the question: How would landmark dynamics affect final airway analysis?

I studied this on the first dynamic data I got from our collaborator. The subject was 59 days old and only had a TC annotation [1]. I then chose the top of the airway as a dummy landmark TVC and the TC for registration. I manually annotated the these two landmarks in each frame and computed the Euclidean difference between landmarks with respect to time.

Figure 4 shows the movement of landmarks between each frame in millimeter. In this figure, TC shows a significant changes after time step 4. The landmark dynamics revealed that the airway of the subject was changing a lot in a period of time.

Figure 5 shows the results of the airway measurement of the 59 days old subject. The difference between the left and right results in Figure 5 is whether considering the landmark dynamics. The left result shows eight functional curves registered with the same TC and TVC annotation at the first frame. The right result shows eight functional curves registered with the updated TC and TVC annotations at each frame. We can see that Figure 5 (right) captured a cross-sectional area drop in the end of the functional data. Figure 5 (left), however, lost this feature (compare the green curves on the left and on the right.) Because an area drop is a critical feature for airway analysis, I considered updating landmarks in each time frame is an important issue and further developed the framework described in Section 2.2.

Since updating landmarks with the subject's deformation is need, I applied the automatically landmark detector described in Section 2. The detector has training stage and detection stage.

Training. The training set had 95 3D CT images with manually annotated landmarks from nasal spine, choana, epiglottis tip, TVC, and TC. I trained a linear SVM classifier for TC using CHOG features. Each subject provided a positive example which would be HOG features extracted from three orthogonal bounding boxes passing through the landmark TC. The scale of the bounding boxes was chosen as twice of the minimum rectangle covering the airway in the axial plane. The negative examples were drawn at an uniform distributed scale and location that did not intersect with the positive bounding boxes. For each positive example, I drawn ten times negative samples.

---

[1]The other annotation nearest TC was TVC, and it was not visible due to the doctor's choice of scan area
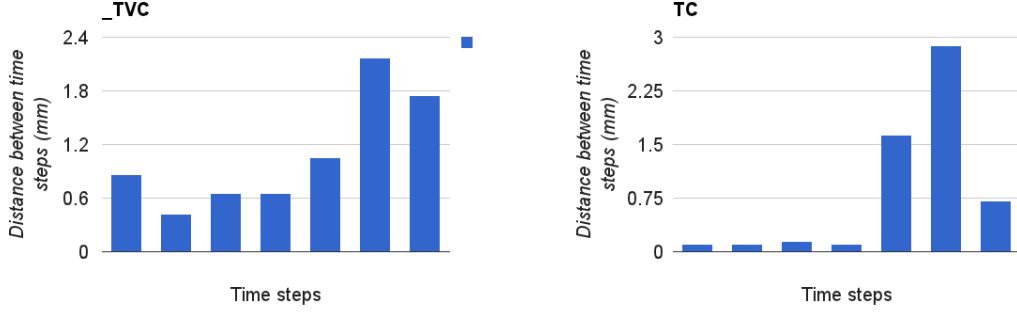
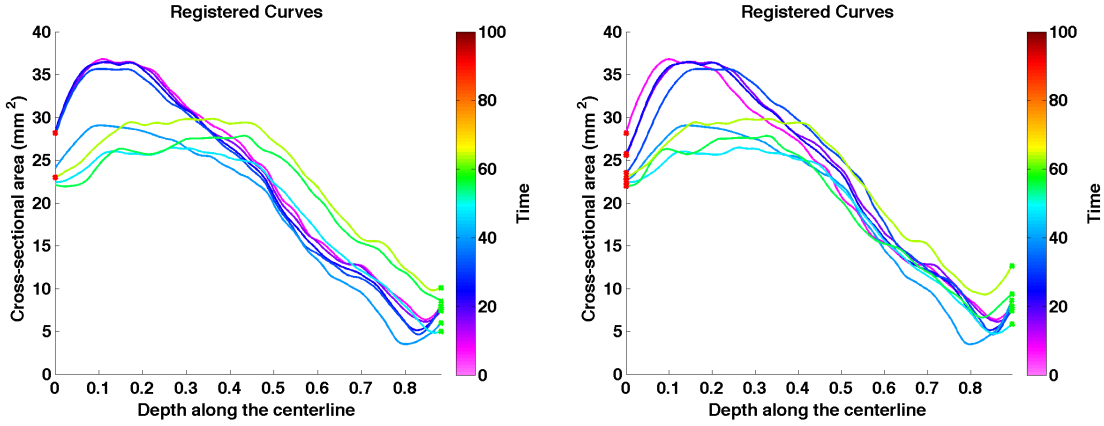Figure 4: Visualization of the landmark dynamics of a 59 days subject.

Figure 5: Cross-sectional area of a 59 days old subject with different registration landmarks. **Left.** The result using the same landmarks for each time step. **Right.** The result with updated landmarks for each time step. Note that in the end of functional data, a lower cross-sectional measurement disappeared on the left but remains on the right.

Detection. For a 512×512×640 3D CT image, there are too many possible hypotheses of scale and location. I utilized a prior of geometry to eliminate possible hypotheses. Once we have the segmentation of the image, we only look at the segmented trachea for landmarks. In each depth I only drew one hypothesis from the center of the airway using the scale with the same heuristic in training. The trained SVM classifier predicted a score of the landmark likelihood given different depth. Figure 6 illustrated the likelihood prediction on training data. We can see there are some outliers in prediction on the training data. Due to these outliers, the mean prediction error was 24.7436 (mm) and the standard deviation was 29.4251 (mm).

Now let us apply the detector on dynamic data and see if it can handle landmark dynamics. Table 1 summarized the landmark dynamics on TC, and my landmark detector using CHOG. The landmark detector works well on the subjects Fleck 004 and Fleck 012. But the detector perform bad on the other subjects.

3.2   Dynamic airway analysis

I performed dynamic airway analysis using functional boxplots on real data provided by a physician. The normal control atlas was built from 68 healthy subjects in control group using the method in [3]. Figure 7 shows the qualitative results. Here I compared these results with the comments from the doctor for two
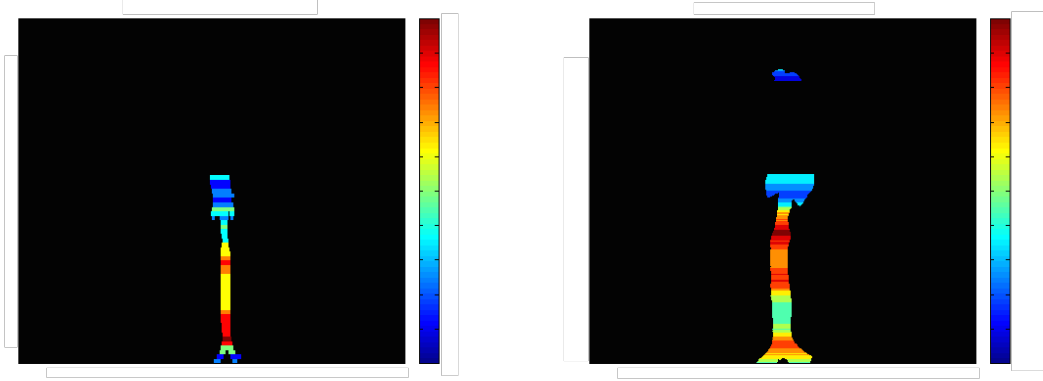
6

Figure 6: Illustration of the likelihood of landmark TC. **Left.** In most cases the detector located TC in a correct position. **Right.** Even failure cases show the correct TC position as their second or third choice.

	Subjects		Landmark Dynamics		Landmark Prediction	
Id	Age (m)	Weight (kg)	mean	std	mean	std
Fleck 004	134.4	39.9	4.75	2.53	3.92	4.46
Fleck 007	114	25.2	4.21	6.16	7.58	5.92
Fleck 008	129.6	48.6	4.48	2.02	23.76	27.83
Fleck 009	130.8	37	1.85	1.35	30.32	3.16
Fleck 010	135.6	43.9	1.93	1.03	21.21	15.34
Fleck 012	42	14.5	1.48	1.03	1.47	0.69
Fleck 013	21.6	16.8	0.89	0.85	23.31	9.06

Table 1: Qualitative measurement of the landmark detection. The prediction errors should be lower than the landmark dynamics. Currently, the detector only met this standard on the subjects Fleck 004 and Fleck 012.

cases.

Fleck-007, 114 months. The subject seemed to have a narrow airway on the CT scan at the first glance. However, comparing with the normal control atlas, this subject definitely had an airway size larger than average normal control data. Also, the doctor's comment was "No significant change in the trachea or bronchi during cough maneuver. No evidence of tracheobronchomalacia.", which qualitatively agrees with the analysis result.

Fleck-008, 129.6 months. For this subject the airway seemed normal in one slice. Yet it was considered narrowed compared with the normal control atlas, and it had large dynamics especially from the depth 0.87 to 1. The depth 0.66 means TVC and 1 means TC. So the subject had narrowed trachea near TC. Quote the doctor's comment "Dynamic, 2.5 cm long narrowing of the mid to distal thoracic trachea caused by the adjacent left-sided aortic arch. The luminal diameter of the trachea changes from 11 mm during diastole to 5 mm in systole", the subject did show problems in the mid to distal thoracic trachea, the analysis results qualitatively agree with the observation.

Table 2 shows qualitative experiments on our dataset. I measured the cross-sectional area changes during the breathing. Studies suggested a subject with over 50% changes of cross-sectional area during the breathing might be a suspected tracheomalacia [17, 18, 19]. Our system computed collapsed dynamics from 4D CT images, and gave qualitative measurement that agree with the studies and the diagnosis from our collaborator.
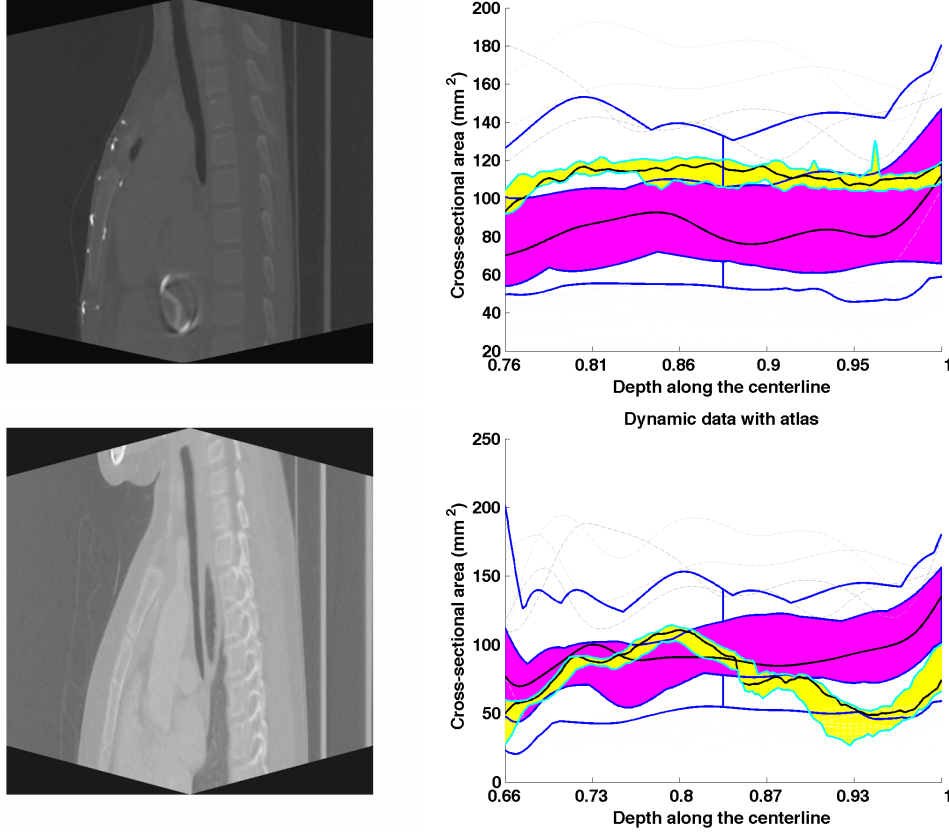
Figure 7: Dynamic airway atlas for subject Fleck-007 and Fleck-008. **Left.** A slice of dynamic data. **Right.** The purple area is the interquartile range of the age-adapted normal control atlas. The blue curves are the maximum and the minimum of the normal control atlas. The yellow area is the entire estimation range of the dynamic subject over the different time steps. The depth is in a unified scale, where 0.66 is TVC and 1 is TC.

4   Discussion

This work proposed an automatic framework for processing a 4D CT image of pediatric airway and providing informative visualization for further analysis. The current approach of data registration using detected landmarks has two issues. First, the detection has outliers. The outliers could be eliminated by examination the distance of each detected result between the detection mean. Yet more powerful features might need to be developed. Second, finding correspondents between a subject without enough landmarks to its statical atlas is handled by heuristics. To fully solving this problem, reducing the rely on landmarks is necessary. Since the within subject deformations were not dramatic, as Guerrero et al.'s suggestion in [5], a deformable image based registration within a subject would be a proper next step.

While the results has showed dynamic ranges and a comparison between dynamic data to normal control atlas, a more straightforward metric which quantifies atypicality of subjects would be very appreciated. A reasonable next approach would be designed a metrics to evaluate the dynamics of airway. A long term goal of this project is to develop a theory of 4D CT atlas building of pediatric airways and aim for going beyond to other parts of human body.

Subjects			Dynamics			Collapsed Dynamics			Diagnosis
Id	Age (m)	Weight (kg)	min	mean	max	min	mean	max	
Fleck 004	134.4	39.9	4%	14%	82%	12%	22%	82%	Normal
Fleck 007	114	25.2	4%	8%	20%	0%	0%	0%	Normal
Fleck 008	129.6	48.6	5%	22%	54%	31%	39%	53%	Tracheomalacia
Fleck 009	130.8	37	7%	15%	43%	0%	0%	0%	Normal
Fleck 010	135.6	43.9	4%	12%	19%	0%	0%	0%	Normal
Fleck 012	42	14.5	11%	18%	44%	0%	0%	0%	Normal
Fleck 013	21.6	16.8	1%	20%	68%	34%	56%	68%	Tracheomalacia

Table 2: Qualitative measurement of atypicality of dynamic subjects. The column Dynamics measures the entire dynamics through the airway. The Collapsed Dynamics measures only on the region which cross-sectional area is smaller than the normal control atlas. Diagnosis is judged by our physician collaborator.

References

[1] Nakano, Y., Whittall, K.P., Kalloger, S.E., Coxson, H.O., Flint, J., Pare, P.D., English, J.C.: Development and validation of human airway analysis algorithm using multidetector row ct. In: Medical Imaging 2002, International Society for Optics and Photonics (2002) 460–469

[2] Abramson, Z., Susarla, S.M., Lawler, M., Bouchard, C., Troulis, M., Kaban, L.B.: Three-dimensional computed tomographic airway analysis of patients with obstructive sleep apnea treated by maxillo-mandibular advancement. Journal of Oral and Maxillofacial Surgery **69**(3) (2011) 677–686

[3] Hong, Y., Davis, B., Marron, J., Kwitt, R., Singh, N., Kimbell, J.S., Pitkin, E., Superfine, R., Davis, S.D., Zdanski, C.J., et al.: Statistical atlas construction via weighted functional boxplots. Medical image analysis **18**(4) (2014) 684–698

[4] Ford, E., Mageras, G., Yorke, E., Ling, C.: Respiration-correlated spiral ct: a method of measuring respiratory-induced anatomic motion for radiation treatment planning. Medical physics **30**(1) (2003)

[5] Guerrero, T., Sanders, K., Castillo, E., Zhang, Y., Bidaut, L., Pan, T., Komaki, R.: Dynamic ventilation imaging from four-dimensional computed tomography. Physics in medicine and biology **51**(4) (2006)

[6] Hill, D.L., Batchelor, P.G., Holden, M., Hawkes, D.J.: Medical image registration. Physics in medicine and biology **46**(3) (2001) R1

[7] Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. Medical Imaging, IEEE Transactions on **32**(7) (2013) 1153–1190

[8] Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. Pattern Analysis and Machine Intelligence, IEEE Transactions on **24**(4) (2002) 509–522

[9] Li, H., Luo, L., Vlasic, D., Peers, P., Popović, J., Pauly, M., Rusinkiewicz, S.: Temporally coherent completion of dynamic shapes. ACM Transactions on Graphics (TOG) **31**(1) (2012) 2

[10] Pizer, S.M., Jung, S., Goswami, D., Vicory, J., Zhao, X., Chaudhuri, R., Damon, J.N., Huckemann, S., Marron, J.: Nested sphere statistics of skeletal models. In: Innovations for Shape Analysis. Springer (2013) 93–115

[11] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Volume 1., IEEE (2005) 886–893

[12] Vondrick, C., Khosla, A., Malisiewicz, T., Torralba, A.: Hoggles: Visualizing object detection features. In: Computer Vision (ICCV), 2013 IEEE International Conference on, IEEE (2013) 1–8

[13] Marron, J., Todd, M.J., Ahn, J.: Distance-weighted discrimination. Journal of the American Statistical Association **102**(480) (2007) 1267–1271

[14] Ramsay, J.O.: Functional data analysis. Wiley Online Library (2006)

[15] Hong, Y., Davis, B., Marron, J., Kwitt, R., Niethammer, M.: Weighted functional boxplot with application to statistical atlas construction. In: MICCAI 2013. Springer (2013) 584–591

[16] Sun, Y., Genton, M.G.: Functional boxplots. Journal of Computational and Graphical Statistics **20**(2) (2011)

[17] Chung, J.H., Kanne, J.P., Gilman, M.D.: Ct of diffuse tracheal diseases. American Journal of Roentgenology **196**(3) (2011) W240–W246

[18] Boiselle, P.M., O'Donnell, C.R., Bankier, A.A., Ernst, A., Millet, M.E., Potemkin, A., Loring, S.H.: Tracheal collapsibility in healthy volunteers during forced expiration: Assessment with multidetector ct 1. Radiology **252**(1) (2009) 255–262

[19] Hasegawa, I., Boiselle, P.M., Raptopoulos, V., Hatabu, H.: Tracheomalacia incidentally detected on ct pulmonary angiography of patients with suspected pulmonary embolism. American Journal of Roentgenology **181**(6) (2003) 1505–1509