

AIRWAY LABELING USING A HIDDEN MARKOV TREE MODEL

James C. Ross^{1,2,3}, Alejandro A. Díaz⁴, Yuka Okajima⁵, Demian Wassermann², George R. Washko⁴,
Jennifer Dy⁶, Raúl San José Estépar^{2,3,5}

¹ Channing Laboratory*, ² Surgical Planning Lab*, ³ Laboratory of Mathematics in Imaging*,
⁴ Pulmonary and Critical Care Division*, ⁵ Department of Radiology*
*Brigham and Women's Hospital, Boston, MA
⁶ ECE Department, Northeastern University, Boston, MA

ABSTRACT

We present a novel airway labeling algorithm based on a Hidden Markov Tree Model (HMTM). We obtain a collection of discrete points along the segmented airway tree using particles sampling [1] and establish topology using Kruskal's minimum spanning tree algorithm. Following this, our HMTM algorithm probabilistically assigns labels to each point. While alternative methods label airway branches out to the segmental level, we describe a general method and demonstrate its performance out to the subsegmental level (two generations further than previously published approaches). We present results on a collection of 25 computed tomography (CT) datasets taken from a Chronic Obstructive Pulmonary Disease (COPD) study.

1. INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is defined as incompletely reversible expiratory airflow obstruction due to emphysematous destruction of the lung parenchyma and remodeling of the small airways [2]; it is now the third leading cause of death in the US [3]. Therefore, it represents a major health concern, and there are ongoing efforts to better understand this complicated disease.

Recent studies have challenged traditional definitions of the disease and suggest connections between the two basic components of COPD: chronic bronchitis (airway disease) and emphysema (lung tissue destruction). For example, the National Heart, Lung, and Blood Institute defines emphysema as "a condition of the lung characterized by abnormal, permanent enlargement of airspaces distal to the terminal bronchiole, accompanied by the destruction of their walls, and without obvious fibrosis" [4]. However, [5] present results suggesting that the narrowing and destruction of terminal bronchioles may precede the loss of acini, thus implicating de-

struction of small airways as possibly causative of emphysema onset.

The authors in [6] reported an association between emphysematous destruction and reduced total airway count as measured by the sum of sixth to eighth generation airways manually determined on volumetric computed tomography (CT), further illuminating the link between emphysema and airway disease. This study indicates that CT can be a valuable tool for investigating the relationship between distal airway disease and emphysema progression and motivates the development of algorithms to automatically quantify the number of airway generations visible on CT.

Anatomically, the first several generations of the human airway tree exhibit a relatively similar topology across subjects, but the topology is known to vary significantly from person to person for more distal branches. To date there have been a number of approaches to assign anatomical names to airway tree branches [7–10]. These approaches limit labeling up to the segmental level (we refer the reader to [11] for the airway labeling scheme adopted here).

Motivated by the need to better explore more distal regions of the airway tree and the usefulness of identifying more distal branches by generation (as opposed to their anatomical labels per se), we propose a novel airway labeling algorithm which assigns specific anatomical names to proximal branches, and labels distal branches according to their branching level: segmental, subsegmental, and subsubsegmental. Our approach is based on Hidden Markov Tree Model (HMTM) analysis applied to discrete samples along the airway tree. We begin by applying particles sampling [1] to acquire the samples. After applying Kruskal's minimum spanning tree algorithm [12] to establish topology on the particles, we invoke the HMTM algorithm described in this paper. In Section II we describe the details of our approach. These include the HMTM representation, and constituent emission probabilities, transition probabilities, and extensions to the Viterbi algorithm for our particular generation labeling task. In Section III we demonstrate the performance of our algorithm, and we draw conclusions in Section IV.

Support provided by NIH grants 1R01HL116931-01, 2R01HL089897-06A1, 2R01HL089856-06A1, K25 HL104085-04, K23HL089353-05, 1P50HL107192, R01HL116473, 1K01 HL118714-01, and R01HL107246.

2. METHODS

For this effort we assume as given a set of samples along the airway tree in the form of scale-space particles [1]. Scale-space particles provide a powerful method for sampling low-level image features of interest, in our case dark tubes (airways) and enable the implicit sampling of airway tree centerlines. Each particle is characterized by its spatial location, orientation, and the scale at which the Hessian response is strongest. Our goal is to assign labels to each of these particles.

The airway tree can be modeled as a directed, acyclic graph; hence, the notion of sequential data naturally arises. We represent the particles data with a graph structure in which nodes represent particles and edges indicate connections between neighboring particles. The resulting graph is undirected and will in general be disconnected. We apply Kruskal’s minimum spanning tree algorithm to the particles point set to build a connected tree [12]. For each subgraph in the spanning tree, each leaf node is considered in turn and tested as a root node candidate. This induces directionality through the graph (from leaves to root) and permits the HMTM labeling described below. The most probable set of airway label assignments is then chosen.

It is important to note that simply relying on branch points in the tree structure as cues for generation changes does not work in general: some branches may be missed, noise branches may be detected, and subgraphs disconnected from the main airway tree don’t have a natural root node. We therefore prefer the probabilistic approach described in this paper. In the following sections we describe the key elements of our main contribution: namely, the HMTM framework for airway labeling. These include HMTM representation, emission and transition probability modeling, and our extensions to the Viterbi algorithm for optimal label assignment.

Hidden Markov Tree Model Representation. We use a first order HMTM to infer labels (the hidden variables) for our particles data. Typical application of hidden Markov models (HMMs) involves a single sequence of observations and associated latent variables [13]. The graphical representation of an example sequence is illustrated in the left of Fig. 1, and the joint distribution over a general sequence of particles (observations) and labels (latent variables) is given by

$$p(\rho^{(1)}, \dots, \rho^{(N)}, \mathbf{g}^{(1)}, \dots, \mathbf{g}^{(N)}) = p(\mathbf{g}^{(1)}) \left[\prod_{n=2}^N p(\mathbf{g}^{(n)} | \mathbf{g}^{(n-1)}) \right] \prod_{n=1}^N p(\rho^{(n)} | \mathbf{g}^{(n)}) \quad (1)$$

where N indicates the number of particles, $\rho^{(n)}$ represents the particle data at sequence point n , and $\mathbf{g}^{(n)}$ represents the corresponding latent variable indicating the airway label.

While the standard HMM assumes that the hidden states follow a linear chain, the hidden latent structure in our case is a tree. The graphical representation of a tree structure is given

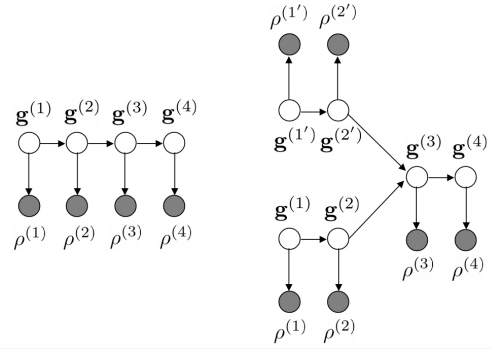


Fig. 1. Left: graphical representation of the sequence of particles $(\rho^{(1)}, \rho^{(2)}, \rho^{(3)}, \rho^{(4)})$ and their associated latent variables $(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \mathbf{g}^{(3)}, \mathbf{g}^{(4)})$. Right: graphical representation of a tree structure. An HMTM can be applied in this instance by considering two sequences: 1) the particles $(\rho^{(1)}, \rho^{(2)}, \rho^{(3)}, \rho^{(4)})$ and associated latent variables $(\mathbf{g}^{(1)}, \mathbf{g}^{(2)}, \mathbf{g}^{(3)}, \mathbf{g}^{(4)})$, and 2) the particles $(\rho^{(1')}, \rho^{(2')}, \rho^{(3)}, \rho^{(4)})$ and associated latent variables $(\mathbf{g}^{(1')}, \mathbf{g}^{(2')}, \mathbf{g}^{(3)}, \mathbf{g}^{(4)})$. Shaded circles represent observed variables; unshaded circles represent unobserved variables.

in the right of Fig. 1, and the general expression for the joint distribution is given by

$$p(\mathcal{P}_i, \mathcal{G}_i) = \prod_{\mathbf{g}^{(l)} \in \mathcal{G}_{leaf}} p(\mathbf{g}^{(l)}) \prod_{n=1}^{|\mathcal{P}_i|} p(\rho^{(n)} | \mathbf{g}^{(n)}) \prod_{\mathbf{g}^{(m)} \in \mathcal{G}_i \setminus \mathcal{G}_{leaf}} p(\mathbf{g}^{(m)} | \text{pa}(\mathbf{g}^{(m)})) \quad (2)$$

where we have used the $|\cdot|$ operator to indicate set cardinality and $\text{pa}(\mathbf{g})$ to represent the parents of \mathbf{g} . \mathcal{P}_i is the set of particles derived from the test image, and \mathcal{G}_i is the set of associated latent variables. The set \mathcal{G}_{leaf} comprises all latent variables corresponding to leaf particles (not including the root). We describe the emission probabilities $p(\rho^{(n)} | \mathbf{g}^{(n)})$ and transition probabilities $p(\mathbf{g}^{(m)} | \text{pa}(\mathbf{g}^{(m)}))$ below.

Emission Probabilities. The emission probabilities give the probability of observing a particle $\rho^{(n)}$ given latent variable $\mathbf{g}^{(n)}$. The observations in our HMTM framework consist of a particle’s spatial location ($\rho_p^{(n)}$), scale ($\rho_s^{(n)}$), and orientation ($\rho_e^{(n)}$). These quantities can be appreciated from Fig. 2.

We propose to use kernel density estimation (KDE) [14] to represent the conditional probability of observing $\rho^{(n)}$ given $g^{(n,i)}$, where $g^{(n,i)}$ represents the i^{th} component of $\mathbf{g}^{(n)}$ ($g^{(n,i)} \in \{0, 1\}$, where $g^{(n,i)} = 1$ indicates the n^{th} particle belongs to label i). KDE is a form of nonparametric density estimation and is a suitable choice in our case given that there is not a clear parametric model for the distribution of the airway labels. To apply KDE we deformably register a collection of labeled particles datasets to the test dataset. The complete set of labeled and registered particles is designated as \mathcal{P}_t , and the subset of particles having latent variable state $g^{(i)}$ is given by $\mathcal{P}_{g^{(i)}}$. Making the assumption that the scale, orientation, and spatial location terms are independent, we



Fig. 2. Close-up of particle glyphs overlaid on an axial CT slice. The glyphs clearly indicate the observed quantities in the HMTM, namely the spatial location, orientation, and scale of each particle (image from [1]).

can express the emission probabilities as

$$p(\rho^{(n)}|g^{(n,i)}) = \frac{1}{N_{g^{(i)}}} \sum_{\rho \in \mathcal{P}_{g^{(i)}}} \mathcal{N}(\rho_s^{(n)} - \rho_s | 0, \sigma_s^2) \times \text{Expon}(\rho_p^{(n)} - \rho_p | \lambda_p) \times \text{Expon}(\angle(\rho_e^{(n)}, \rho_e) | \lambda_\angle) \quad (3)$$

where $N_{g^{(i)}}$ is the number of particles labeled as $g^{(i)}$, $\mathcal{N}(\cdot | \mu, \sigma^2)$ represents a normal distribution with mean μ and variance σ^2 , and $\text{Expon}(\cdot | \lambda)$ represents the exponential distribution with rate parameter λ . The parameters σ_s^2 , λ_p , and λ_\angle are learned from training data.

Transition Probabilities. Transition probabilities indicate the probability of transitioning from one latent state to another and are captured in a transition matrix. The transition matrix in our model, \mathbf{A} , is a function of both the change in scale and the change in direction between a parent particle and its child:

$$\mathbf{A}_{ij}(\rho^{(p)}, \rho^{(n)}) \sim p(\rho_s^{(p)} - \rho_s^{(n)}, \angle(\rho_e^{(p)}, \rho_e^{(n)}) | i \rightarrow j) \times p(i \rightarrow j) \quad (4)$$

where $\rho^{(p)}$ is a parent of $\rho^{(n)}$, and $i \rightarrow j$ indicates the transition from state i to state j . The likelihood and prior terms given in Eq. 4 are learned from the labeled training data set.

With the elements of \mathbf{A} defined, we can now express the conditional probability of $\mathbf{g}^{(n)}$ given its parents as

$$p(\mathbf{g}^{(n)} | \text{pa}(\mathbf{g}^{(n)})) = \frac{1}{Z(\text{pa}(\mathbf{g}^{(n)}))} \prod_{\mathbf{g}^{(p)} \in \text{pa}(\mathbf{g}^{(n)})} \left[\prod_{j=1}^{N_g} \prod_{k=1}^{N_g} A_{jk}^{g^{(p,j)} g^{(n,k)}} \right] \quad (5)$$

where N_g is the number of states, and $Z(\text{pa}(\mathbf{g}^{(n)}))$ is a normalization constant given by

$$Z(\text{pa}(\mathbf{g}^{(n)})) = \sum_{k=1}^{N_g} \left[\prod_{\mathbf{g}^{(p)} \in \text{pa}(\mathbf{g}^{(n)})} \prod_{j=1}^{N_g} A_{jk}^{g^{(p,j)} g^{(n,k)}} \right] \quad (6)$$

Extending the Viterbi Algorithm. The Viterbi algorithm is typically used to find the most probable sequence of latent

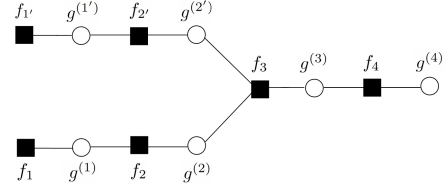


Fig. 3. Factor graph representation of the graphical model shown in the right of Fig. 1.

variable states for a given observation sequence [15]. In order to extend the algorithm to the tree structures that characterize our data, it is helpful to consider a factor graph representation of our directed graphical model. Factor graphs rely on the fact that directed (and undirected) graphs allow the corresponding joint distribution to be defined as a product of factors, f , over subsets of variables [16]. Factor graphs can be generated from directed graphs by inserting additional nodes between variable nodes for each of the factors. Referring again to the graphical model shown in the right of Fig. 1, the corresponding factor graph representation is shown in Fig. 3, and the general expression for the factors for $n > 1$ is given by

$$f_n(\text{pa}(\mathbf{g}^{(n)}), \mathbf{g}^{(n)}) = p(\rho^{(n)} | \mathbf{g}^{(n)}) p(\mathbf{g}^{(n)} | \text{pa}(\mathbf{g}^{(n)})) \quad (7)$$

and for $n = 1$

$$f_1(\mathbf{g}^{(1)}) = p(\mathbf{g}^{(1)}) p(\rho^{(1)} | \mathbf{g}^{(1)}) \quad (8)$$

Note that we have adopted a simplified factor graph representation in which the observation variables, ρ , are not explicitly represented, and the emission probabilities have been combined with the transition probabilities in our expression for the factors (Eq. 7).

The Viterbi algorithm can be seen as a specific application of the max sum algorithm applied to HMMs [14]. The max sum algorithm can be realized with a message passing framework using the factor graph representation. The algorithm proceeds with a forward recursion stage involving factor-to-variable and variable-to-factor messages followed by a backtracking stage that specifically identifies latent variable values that maximize the joint distribution. In our case the factor-to-variable message is given by

$$\mu_{f_{n+1} \rightarrow \mathbf{g}^{(n+1)}}(\mathbf{g}^{(n+1)}) = \max_{\text{pa}(\mathbf{g}^{(n+1)})} \left\{ \ln f(\text{pa}(\mathbf{g}^{(n+1)}), \mathbf{g}^{(n+1)}) + \sum_{\mathbf{g}^{(p)} \in \text{pa}(\mathbf{g}^{(n+1)})} \mu_{\mathbf{g}^{(p)} \rightarrow f_{n+1}}(\mathbf{g}^{(p)}) \right\} \quad (9)$$

and the variable-to-factor message is

$$\mu_{\mathbf{g}^{(n)} \rightarrow f_{n+1}}(\mathbf{g}^{(n)}) = \mu_{f_n \rightarrow \mathbf{g}^{(n)}}(\mathbf{g}^{(n)}) \quad (10)$$

Now letting $\omega(\mathbf{g}^{(n)}) = \mu_{f_n \rightarrow \mathbf{g}^{(n)}}(\mathbf{g}^{(n)})$ we have the following recursion expression

$$\omega(\mathbf{g}^{(n)}) = \ln p(\rho^{(n)} | \mathbf{g}^{(n)}) + \max_{\mathbf{g}^{(p)} \in \text{pa}(\mathbf{g}^{(n)})} \left\{ \ln p(\mathbf{g}^{(n)} | \text{pa}(\mathbf{g}^{(n)})) + \sum_{\mathbf{g}^{(p)} \in \text{pa}(\mathbf{g}^{(p)})} \omega(\mathbf{g}^{(p)}) \right\} \quad (11)$$

The algorithm proceeds by performing the forward recursion, updating the $\omega(\mathbf{g}^{(n)})$ terms along the way. During the forward recursion we also maintain a mapping, ψ , from latent variable states to parent latent variable states that maximize contributions to the summation Eq. 11. This mapping enables us to perform backtracking after we have completed the forward recursion, thus specifically defining latent variable states (i.e. airway labels) that are jointly most probable given our observation (particle) data.

3. EXPERIMENTS AND RESULTS

25 particles datasets were manually labeled by two pulmonologists. We performed a series of leave-one-out tests: 24 datasets were used as the atlas set, and the remaining dataset was used as the test set. For each test the parameters described above for the emission and transition probabilities were learned from the training set.

Tables 1 and 2 show quantitative results of our experiments in the form of normalized confusion matrices: each entry indicates the fraction that the algorithm-assigned labels either agree with the reference standard label or are confused with another label; perfect agreement is indicated by a value of 1.0 in each of the shaded entries. The values are computed across all 25 leave-one-out experiments. Results in Table 1 show agreement using KDE classification using emission probabilities only and indicate performance if each particle is labeled irrespective of the overall tree structure to which it belongs. Conversely, the results in Table 2 correspond to the final algorithm output after HMTM inference has been performed over the tree structure as a whole.

There is a noticeable performance increase for the most distal airway branches (segmental, sub-segmental, and subsub-segmental) when using the HMTM model. There is a slight dip in performance for the superior division bronchus (SDB) and lingular bronchus (LB), but it is worthwhile to note that both of these branches are quite short. When considering classification performance overall, the KDE approach has an accuracy of 61.2%, while the complete HMTM model has an improved accuracy of 71.3%.

Another desirable feature of the complete HMTM model over simple KDE classification is that it takes the tree structure into account and prevents impossible state transitions. This can be clearly seen in Fig. 4. The KDE classifier can,

Table 1. Results using KDE classification. T = trachea, MB = main bronchi, ULB = upper lobe bronchus, SDB = superior division bronchus, LB = lingular bronchus, MLB = middle lobe bronchus, IB = intermediate bronchus, LLB = lower lobe bronchus, S = segmental bronchus, SS = sub-segmental bronchus, SSS = subsub-segmental bronchus.

	T	MB	ULB	SDB	LB	MLB	IB	LLB	S	SS	SSS
T	0.99	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MB	0.03	0.91	0.00	0.00	0.00	0.00	0.03	0.02	0.00	0.00	0.00
ULB	0.00	0.00	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SDB	0.00	0.00	0.03	0.90	0.00	0.00	0.00	0.00	0.05	0.02	0.00
LB	0.00	0.01	0.03	0.00	0.85	0.00	0.00	0.00	0.09	0.01	0.01
MLB	0.00	0.00	0.00	0.00	0.00	0.94	0.00	0.00	0.05	0.01	0.00
IB	0.01	0.01	0.00	0.01	0.00	0.00	0.84	0.14	0.00	0.00	0.00
LLB	0.00	0.01	0.00	0.00	0.00	0.00	0.13	0.80	0.05	0.02	0.00
S	0.00	0.00	0.01	0.01	0.03	0.02	0.00	0.06	0.62	0.19	0.08
SS	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.02	0.21	0.44	0.31
SSS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.31	0.58

Table 2. Results using HMTM. Abbreviations are as in Table 1.

	T	MB	ULB	SDB	LB	MLB	IB	LLB	S	SS	SSS
T	0.99	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MB	0.06	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ULB	0.00	0.04	0.95	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
SDB	0.00	0.00	0.06	0.73	0.00	0.00	0.00	0.00	0.21	0.00	0.00
LB	0.00	0.00	0.01	0.00	0.78	0.00	0.00	0.00	0.21	0.00	0.00
MLB	0.00	0.00	0.00	0.00	0.00	0.95	0.00	0.01	0.04	0.00	0.00
IB	0.01	0.04	0.00	0.00	0.00	0.00	0.84	0.10	0.00	0.00	0.00
LLB	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.83	0.11	0.00	0.00
S	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.04	0.67	0.25	0.03
SS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.59	0.28	0.00
SSS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.26	0.71

e.g., label consecutive particles as sub-segmental, subsub-segmental, and then sub-segmental again, which is an impossible progression. The HMTM model prevents such scenarios and enforces only allowable state transitions.

4. CONCLUSION

We have introduced a novel, probabilistic method – Hidden Markov Tree Model – for assigning labels to samples of the airway tree represented by particle points. Previous approaches have focused on labeling out to the segmental level; we show results out to the subsub-segmental level. Our framework is general and can in theory be applied to even more distal branches provided a labeled atlas is given. The framework we present can also be applied to airway trees that are not fully connected and to trees with noise branches and missing branches. Space limitations prevent a full presentation of these results for this paper.

5. REFERENCES

- [1] G. L. Kindlmann, R. San José Estépar, S. M. Smith, and C. F. Westin, “Sampling and visualizing creases with scale-space particles,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 6, pp. 1415–1424, 2009.
- [2] J. Hogg and W. Timens, “The pathology of chronic obstructive pulmonary disease,” *Annu Rev Pathol*, vol. 4, pp. 435–459, 2009.

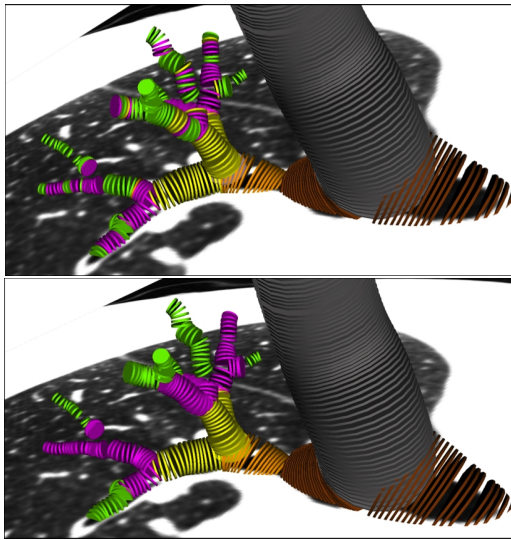


Fig. 4. Top: Results using KDE classification. Bottom: results using HMTM. Branches are color-coded as follows: gray = trachea, brown = main bronchi, orange = upper lobe bronchus, yellow = segmental bronchi, magenta = sub-segmental bronchi, green = subsub-segmental bronchi.

- [3] A. M. Miniño, S. L. Murphy, J. Xu, and K. D. Kochanek, “Deaths: final data for 2008.” *National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System*, vol. 59, no. 10, p. 1, 2011.
- [4] S. GL, K. J, T. WM, and B. ZH, “The definition of emphysema: report of a national heart, lung, and blood institute workshop,” *Am Rev Respir Dis*, vol. 132, pp. 182–5, 1985.
- [5] J. McDonough, R. Yuan, M. Suzuki, and et al, “Small-airway obstruction and emphysema in chronic obstructive pulmonary disease,” *N Engl J Med*, vol. 365, no. 17, pp. 1567–1575, 2011.
- [6] A. Diaz, C. Valim, and T. Y. et al, “Airway count and emphysema assessed by chest ct imaging predicts clinical outcome in smokers,” *Chest*, vol. 138, no. 4, pp. 880–887, 2010.
- [7] A. Feragen, J. Petersen, M. Owen, P. Lo, L. Thomsen, M. Wille, A. Dirksen, and M. de Bruijne, “A hierarchical scheme for geodesic anatomical labeling of airway trees,” in *MICCAI*, vol. 7512, 2012, pp. 147–155.
- [8] P. Lo, E. van Rikxoort, J. Goldin, F. Abtin, M. de Bruijne, and M. Brown, “A bottom-up approach for labeling of human airway trees,” in *MICCAI*, vol. Fourth International Workshop on Pulmonary Image Analysis, 2011, pp. 23–34.
- [9] B. van Ginneken, W. Baggeman, and E. van Rikxoort, “Robust segmentation and anatomical labeling of the airway tree from thoracic ct scans,” in *MICCAI*, vol. 5241, 2008, pp. 219–226.
- [10] J. Tschirren, G. McLennan, K. Palagyi, E. Hoffman, , and M. Sonka, “Matching and anatomical labeling of human airway tree,” *IEEE Transactions on Medical Imaging*, vol. 24, no. 12, pp. 1540–1547, 2005.
- [11] F. H. Netter, *Atlas of human anatomy*. Elsevier Health Sciences, 2010.
- [12] J. Kruskal, “On the shortest spanning subtree of a graph and the traveling salesman problem,” in *Proceedings of the American Mathematical Society*, vol. 7, no. 1, 1956, pp. 48–50.
- [13] L. R. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [14] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2007.
- [15] A. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.
- [16] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.