Automated Quantitative Assessment of Lung Fissure Integrity on CT

Matthew Brown, Robert Ochs, Fereidoun Abtin, Arash Ordookhani, Megan Brown, Hyun Kim, Greg Shaw, Daniel Chong, Jonathan Goldin

Department of Radiological Sciences, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

M Brown – mbrown@mednet.ucla.edu; R Ochs – rochs@ucla.edu; F Abtin – fabtin@mednet.ucla.edu; A Ordookhani – aordookhani@mednet.ucla.edu; M Brown – meganbrown@ mednet.ucla.edu; G Shaw – gshaw@ucla.edu; D Chong – dchong@mednet.ucla.edu; J Goldin – Jgoldin@mednet.ucla.edu

Abstract. This study investigates an unsupervised machine learning approach for quantitative analysis of pulmonary fissure completeness. The analysis of pulmonary fissures has traditionally been a subjective task, relying on visual assessment. Fissure analysis is becoming an important consideration as clinical studies suggest that fissural completeness significantly correlates with success of new endobronchial valve therapies for emphysema. Using an unsupervised clustering approach and 600 fissure samples, three clusters emerged: fissure, non-fissure, and bronchovascular bundle. The performance of the system was then evaluated using 84 randomly selected fissure images. The test cases were also independently contoured by two observers to form a reference standard for fissural completeness. Analysis of the results showed that there was no statistically significant difference between the CAD system and the human observers in calculating fissure percentage completeness (T-Test $P > 0.05$). Pair-wise comparisons of CAD-reader and reader-reader assessment of fissure completeness showed comparable levels of agreement ≥77%.

Keywords: CAD, Fissure, Emphysema,

1 Introduction

Quantitative Image Analysis (QIA) is an important component in the process of patient selection and treatment targeting in new emphysema treatments involving minimally invasive lung volume reduction. One of the new treatments utilizes one-way endobronchial valves placed in the segmental airways to exclude and deflate an emphysematous lobe without the need for surgery, thereby allowing other (healthier) lobes to further expand and improve lung function [1].

The degree of fissure integrity (i.e., completeness) is emerging as a potential predictor of treatment efficacy (deflation of emphysematous lobe) [2]. An incomplete pulmonary fissure, as shown in Figure 1, indicates the potential for collateral ventilation between adjacent lobes, circumventing complete lobar occlusion and lobar collapse [3,4]. Thus analysis of fissure completeness may play a key role in

identifying patients who would benefit from the therapy and determining the target lobe for the endobronchial valve treatment.

Determination of fissural completeness using HRCT imaging has previously been carried out by visual inspection and subjective grading into three categories: complete, incomplete, and absent [2,5,6]. This is a tedious and difficult task requiring review of a large number of images under multi-planar reformats, which has motivated the development of an automated quantitative system to assess fissure integrity. A variety of methods have been introduced for fissure detection [7-13], but none have been used to quantify fissural completeness or proven successful in an emphysematous lung [14]. In emphysema patients, the lung parenchyma is inconsistent in structure due to the presence of enlarged air-sacs known as bullae which lead to fissures that are irregular and indistinct on imaging.

An automated quantitative system could be of tremendous benefit in making assessment of fissure integrity more discriminative, reproducible, and broadly applicable. The aim of this pilot study is to demonstrate the feasibility of automated quantitative assessment of fissure integrity from thin-section CT in emphysema subjects with abnormal and incomplete pulmonary fissures.

2 Methods

2.1 Image Data Collection

The CT images used in this study were selected at random from a cohort of 486 emphysema subjects from a research database (see cohort demographics in Table 1). Images were acquired with the following imaging parameters: 120 kVp, 140 to 300 mAs, and a pitch ranging from 0.984 to 1.5. Images were reconstructed with slice thicknesses ranging from 1 to 3 mm and using standard reconstruction algorithms (e.g., GE STD, Siemens B30f, Philips B, and Toshiba FC10 filters).

For this pilot study we focused our initial analyses on the left major fissure. The research database contains 12,391 technologist-drawn and radiologist approved 2D left major fissure contours from 486 subjects. From these contours samples were randomly selected to form independent training and test sets.

Table 1. Demographics of 486 emphysema subjects in a research database.

	Absolute (Mean ± SD)	Range
Age (y)	63 ± 7	41 – 76
PFT TLC (L)	7.62 ± 1.48	3.4 – 12.14
PFT RV (L)	4.83 ± 1.21	0.86 – 9.22
CT TLC (L)	6.99 ± 1.38	3.98 – 10.45
CT RV (L)	5.12 ± 1.26	2.29 – 8.99
Voxels below -910 HU (%)	56.53± 10.30	30.45 – 81.27

2.2 Image Features

The Hessian matrix of second order partial derivatives (Eq. 1) was used to characterize the variations in intensity about a point. In our initial implementation the Hessian matrix was computed in two-dimensions, however, the technique can be generalized to three-dimensions [13]. The computation has been described in greater detail by other authors [7,15]. The eigenvalues of the Hessian matrix, κ_1, κ_2, are computed and ranked according to their absolute value $|\kappa_1| > |\kappa_2|$. Prior to computation of the Hessian, Gaussian smoothing was applied to the image with a kernel with standard deviation σ = 1.0 mm.

$$Hf(x,y) = \begin{bmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy} \end{bmatrix}. \quad \textbf{(1)}$$

Fissures are modeled as faint, plate-like structures (or lines in 2D images) due to their thin surface and partial volume averaging. κ_1 should correspond to the gradient change normal to the fissural line (along the first eigenvector). The other orthogonal eigenvalue (κ_2) should correspond to the second eigenvector in the direction of the fissure and should be closer to zero. Based on this assumption we compute a "plateness image" using Equation 2 (see Figure 1).

$$I = 1 - \left| \frac{\kappa_2}{\kappa_1} \right| \text{, where the ratio } \kappa_2 / \kappa_1 \text{ is expected to be high for fissures.} \quad \textbf{(2)}$$

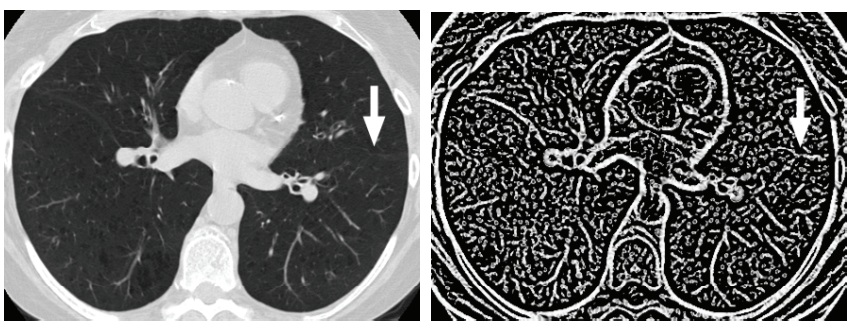


Figure 1. Original (on left) and plateness (on right) images with an incomplete left major fissure indicated by arrows. The fissures appear as faint white lines in the original image and solid lines in the plateness image. Since there are blood vessels with similar appearance the left major fissure is marked with an arrow.

Three features are computed for each voxel along the fissure path within a 32 x 32 patch around the voxel from the original CT and plateness images as shown in Table 2. The fissure path is based on a manually drawn contour as described in section 2.1.

Table 2. Features computed along fissure path in a 32 x 32 patch around the voxel of interest in the original CT and plateness (gradient-based) images.

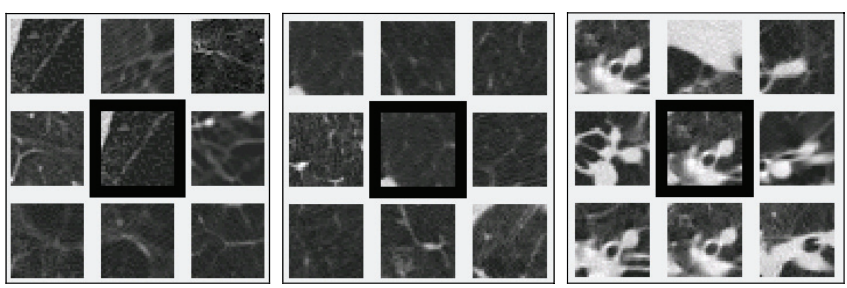
Image	Feature	Range	Expected values
Original CT	Median gray level	0 - 255	higher if voxel is a fissure
Plateness	20th centile of histogram	0 - 255	higher if voxel is a fissure
Plateness	Max run length of voxels with intensity > 100	0 - 32	longer if voxel is a fissure

2.3 Classification Model

The classification model was machine-learned using unsupervised k-means clustering. The number of clusters was set as K=3 based on the premise that three classes are expected: fissure, non-fissure, bronchovascular bundle (near the hilum).

The classifier was trained using 600 feature samples from among the 12,391 fissure contours in our research database. Each sample (computed from a 32x32 patch) was taken from a different fissure contour (i.e., different CT image) to minimize dependence between the samples and bias in the classifier. Sample patches surrounding the learned cluster centers are shown in Figure 2.

Using the cluster centers, a minimum-distance classifier was implemented to classify each voxel along the fissural path. For each voxel along the path, the three features are computed in the 32x32 neighborhood and input to the classifier. The fissural completeness is then calculated as the percentage of pixels along the path classified as fissure.



(a) (b) (c)

Figure 2. Nine training samples nearest to the learned cluster centers for (a) fissures, (b) non-fissures, and (c) bronchovascular bundles.

2.4 Experimental Testing

100 HRCT slices were selected at random from our research database. Two human observers independently reviewed all 100 slices. They were shown the endpoints of the fissure marked initially by the technologist (which was also used as the input fissure path to the CAD system). If either one of the observers in the current

experiment disagreed with the endpoints of these previous contours the case was rejected. 16 cases were rejected during this process, resulting in 84 CT images which the observers and CAD system independently analyzed.

The human observers were trained on lung anatomy, software application, and sensitivity calibration. We did not use radiologists for this pilot study due to the previous determination that well-trained observers produce similar inter-observer variability as licensed radiologists [16].

The observers assessed the entire fissural path within each CT image, they marked portions of the fissure they considered to be intact and left blank the parts of the fissure they considered to be absent. They also marked portions of the path passing through vessel bundles (near the hilum) which were excluded from completeness calculations. The CAD system was also applied to each case and automatically classified voxels along the fissural path into the same three categories. For each observer the fissure integrity was then assessed by calculating the percent-completeness of the fissure:

$$Completeness = \frac{N_{Fiss}}{N_{Fiss} + N_{NonFiss}} \times 100 \ . \qquad \textbf{(3)}$$

where N_{Fiss} = number of voxels classified as fissure and $N_{NonFiss}$ = number of voxels classified as non-fissure.

The completeness percentage values were then compared pair-wise between Reader 1, Reader 2, and CAD. The percentages were also converted into a binary decision of complete or incomplete for each observation. Based on previous studies involving visual assessment, a completeness percentage of ≥ 90% was considered as a complete fissure for potential treatment planning [16]. Bland-Altman plots (Figure 5) were used to quantify statistically significant agreements between treatments. Paired T tests were also performed to further analyze the agreement between treatments (See Section 3).

3. Results

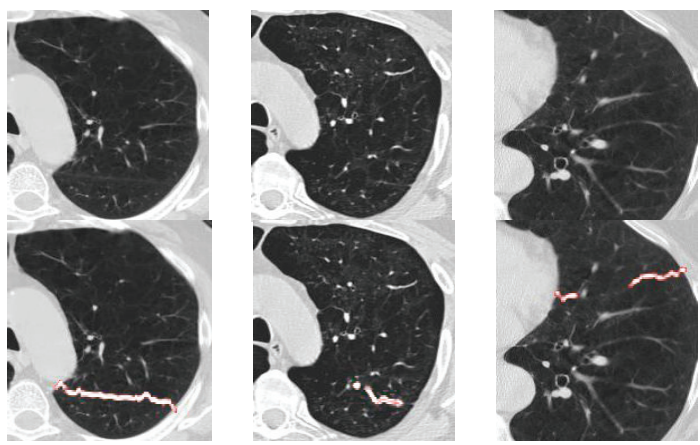
Figures 4a-c are scatter plots showing the pair-wise comparisons between Reader 1, Reader 2, and CAD. Figure 4a compares the two readers, and Figures 4b and 4c compare CAD against each of the readers. Dashed lines indicate 90% fissural completeness. Points above/right of these lines (in the upper right quadrant) represent agreement on complete fissures, while points in the lower left represent agreement on incomplete fissures. The upper left and lower right quadrants represent disagreement between the observers. The numbers, *N*, shown in each quadrant are the number of pair-wise observations in each category. For example, in Figure 4a there were 55 fissures where both readers agreed that the fissure was complete, 18 where they agreed it was incomplete and 6+5=11 where they disagreed. Similar agreement/disagreement counts are shown for Figures 4b and 4c involving CAD vs Reader.

Bland-Altman plots were also used to determine the pair-wise agreement between fissure completeness scores. All three plots contain a cluster of points with zero difference at 100% completeness. In each of the Bland-Altman plots, 6 readings were significantly different (with 95% confidence), leaving a 78-reading agreement for each comparison. The paired T tests resulted in p values of 0.3401, 0.0863, and 0.3876 for the comparisons between Reader 1 and Reader 2, Reader 1 and CAD, and Reader 2 and CAD respectively.

Table 3 summarizes the agreement between the three observers. The first row indicates that all three agree on completeness/incompleteness in 72.6% of the fissures. In 14.3% of the fissure the two readers agree but CAD does not. In the remaining cases CAD agrees with one of the readers. Figure 3 shows examples of fissures with different levels of fissure completeness and observer agreement.

Table 3. Agreement counts between Reader 1, Reader 2, and the CAD system.

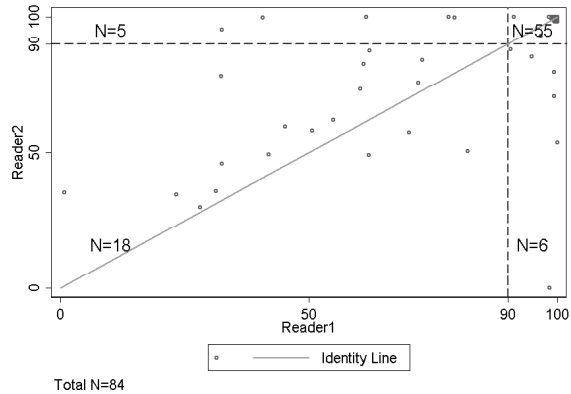
Agreement On Fissural Completeness			
Readers that agree	**≥90%**	**<90%**	**Total**
R1, R2, and CAD	49	12	61 (72.6%)
Only R1 and R2	6	6	12 (14.3%)
Only R1 and CAD	5	2	7 (8.3%)
Only R2 and CAD	3	1	4 (4.8%)



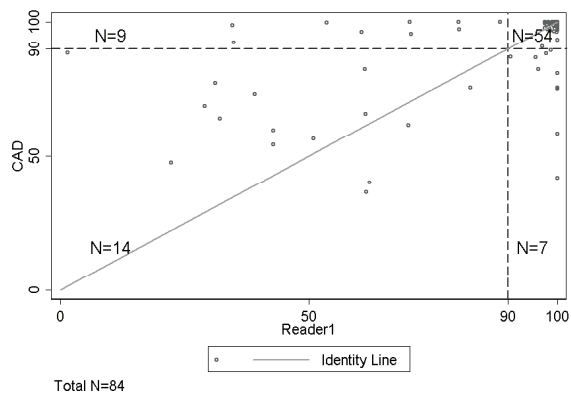
(a) (b) (c)

Figure 3. Original CT images and CAD detected fissures for cases where (a) both readers and CAD classified the fissure as complete, (b) both readers and CAD classified the fissure as incomplete, and (c) one reader and CAD classified the fissure as incomplete and the other reader classified it as complete.

(a)



(b)



(c)

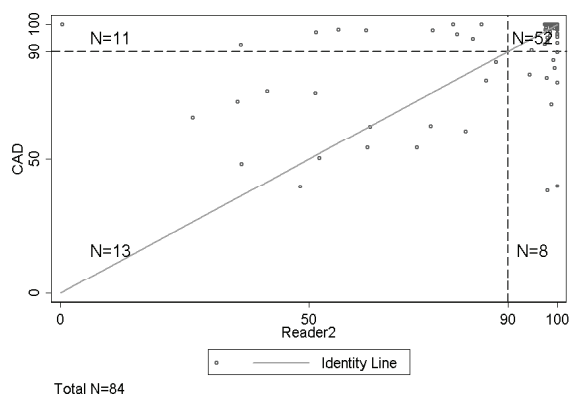


Figure 4. Scatter plot of fissure percentage completeness for (a) Reader 2 vs Reader 1, (b) CAD vs Reader 1, and (c) CAD vs Reader 2.

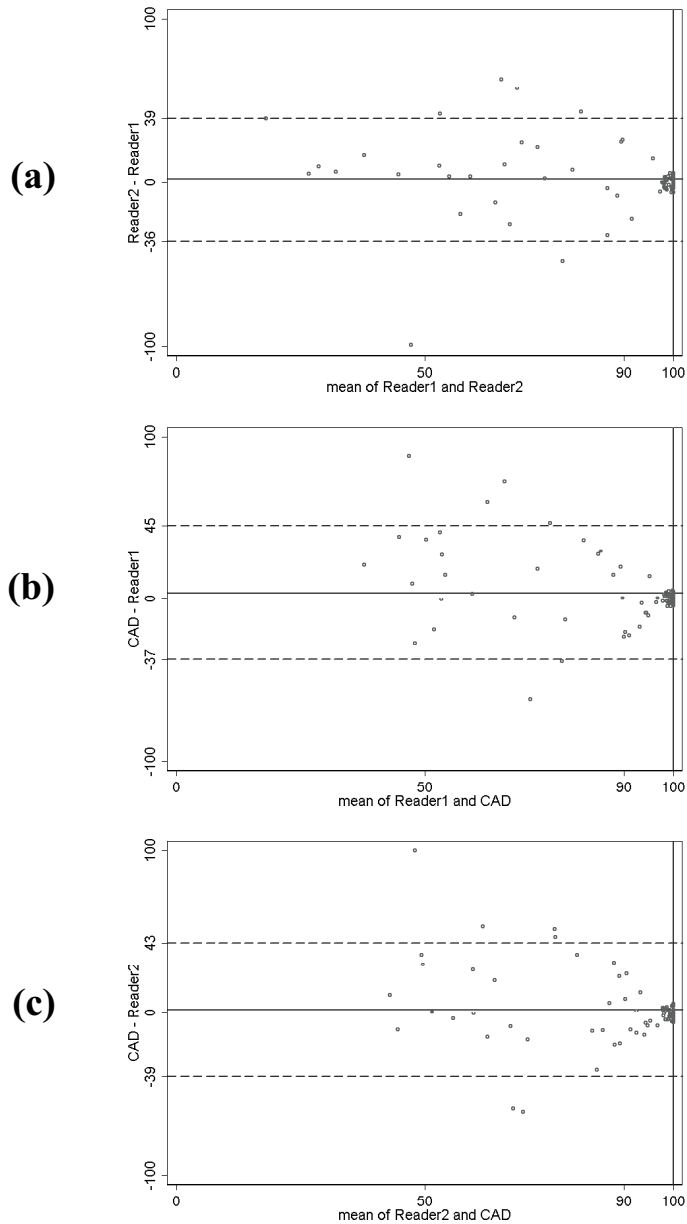


Figure 5. Bland-Altman plot of fissure completeness percentages for (a) Readers 1 and 2, (b) CAD and Reader 1, and (c) CAD and Reader 2.

4 Discussion

There was good pair-wise agreement for the percentage of fissural completeness, with no statistically significant differences within any of the three comparisons. When applying the 90% threshold for completeness CAD had good agreement with each of the readers (81% and 77% respectively - see Figures 4b,c), although the agreement between the readers themselves was higher (88% - see Figure 4a). This performance is particularly encouraging since it involved subjects with emphysema and incomplete fissures, rather than clearly defined normal fissures as has typically been the case in previous work on fissure detection. The difficultly of the fissure classification problem in emphysema subjects is indicated by the agreement levels between the readers. The scatter plot in Figure 4a shows that even trained readers have quite different completeness scores in some (difficult) cases.

We developed our prototype system for the left major fissure since its appearance seems more consistent than the right major or minor fissures. The completed system will include fissures from the right lung which may necessitate a larger sample size for classifier training. In this pilot study we used images with a range of CT technical factors, and further systematic investigation on their influence on features is needed. We will also continue investigation of supervised machine-learning approaches with expert-labeled voxels [13].

In future studies we expect improved agreement. For this pilot study we intentionally focused on a small number of features that were selected a priori and a simple classification method. Thus we find the results to be very encouraging and expect the approach to generalize well to larger data sets. Also, we were using only one CT image per assessment, both in terms of CAD feature calculation and for reader review. This was done for speed/simplicity of calculation and observer review in this initial investigation of the feasibility of using CAD to assess abnormal fissures. Previous fissure detection approaches have used multiple slices (three dimensions) and lung vasculature data [7-13] to guide fissure detection. We will extend the features to 3D and the reader evaluation to multiple slices and expect to achieve better performance and greater agreement. We will also investigate the feature space further with a labeled training set to determine whether a non-linear classifier is appropriate.

By maintaining a high accuracy for detection of complete and almost complete fissures we aim to maximize the number of successful outcomes for the endobronchial valve treatment [1,2].

5 Conclusion

The CAD prototype system showed good agreement with human observers in computing the percentage of fissure completeness and in classifying the fissures as complete or incomplete. The pair-wise agreement between CAD and each reader was comparable to that between the readers themselves. This work is one of the first efforts to specifically detect abnormal/incomplete fissures in subjects with emphysema and has important clinical applications in targeting of endobronchial valve treatments. The results are very encouraging for a challenging problem and we

expect the extension of the features to three-dimensions will yield further improvements in performance.

References

1. Strange, C., Herth, F., Kovitz, K.L., McLennan, G., Ernst, A., Goldin, J., Noppen, M., Criner, G.J., Sciurba, F.C.: Design of the Endobronchial Valve for Emphysema Palliation Trial (VENT): a non-surgical method of lung volume reduction. BMC Pulm Med 2007; 7(1):10 (2007)
2. Abtin, F., Goldin, G.G., Strange, C., Criner, G.J., Marquette, C., Sciurba, F.C., Brown, M.S., Rao, A., Kim, H., Roback, D., Jiang, A., Pais, R., Irani, Z., Herth, F.J., Ernst, A., Kovitz, F.L., McLennan, G.: The Influence of Fissural Anatomy on the Treatment Outcome of Patients with Emphysema. Posterboard presentation #34;A755 at ATS at Toronto (2008)
3. Scanlon, T.S., Benumof, J.L.: Demonstration of interlobar collateral ventilation. J Appl Physiol 46:658--661 (1979)
4. Hogg, J.C., Macklem, PT, Thurlbeck, W.M.: The resistance of collateral channels in excised human lungs. J Clin Invest. 48:421--431 (1969)
5. Proto AV, Ball JB Jr. Computed tomography of the major and minor fissures. Am J Roentgenol 140:439–448 (1983)
6. Ariyurek, O.M., Karabulut, N., Yelgec, N.S., Gulsun, M.: Anatomy of the minor fissure: assessment with high-resolution CT and classification. Eur Radiol. 12:175-180 (2002)
7. Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., and Kikinis, R.: Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. Med. Image Anal. 2(2):143--168 (1998)
8. Zhang L., Hoffman, E.A., Reinhardt, J.M.: Lung lobe segmentation by graph search with 3-D shape constraints, Proc. SPIE. 4321:204--215 (2001)
9. L. Zhang, et al., Atlas-driven lung lobe segmentation in volumetric X-ray CT images, Proc. SPIE. 5031:306--315 (2003)
10. Wang, J., Betke, M., Ko, J.P., Shape-based curve growing model and adaptive regularization for pulmonary fissure segmentation in CT, in Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, MICCAI. 3216: 541-548 (2003)
11. Saita S, Yasutomo M, Kubo M, Kawata Y, Niki N, Eguchi K, Ohmatsu H, Kakinuma R, Kaneko M, Kusumoto M, Moriyama N, Sasagawa M. An extraction algorithm of pulmonary fissures from mutli-slice CT image. Proc SPIE. 5370:1590--1597 (2004)
12. van Rikxoort EM, van Ginneken B, Klik M, Prokop M. Supervised enhancement filters: application to fissure detection in chest CT scans. IEEE Trans Med Imaging. 2008 Jan;27(1):1-10.
13. Ochs R., Goldin J., Abtin F., Kim H.J., Brown K., Batra P., Roback D., McNitt-Gray M., M. Brown. Automated classification of lung bronchovascular anatomy in CT using AdaBoost. Medical Image Analysis Volume 11, Issue 3, June 2007, Pages 315-324
14. Sluimer, I., Schilham, A., Prokop, M., van Ginneken, B.: Computer analysis of computed tomography scans of the lung: a survey. IEEE Trans. Med. Imaging 4, 385--405 (2006)
15. Krissian, K., Malandain, G., Ayache, N., Vaillant, R., Trousset, Y. Model Based Detection of Tubular Structures in 3D Images. Computer Vision Image Understanding 80, 130-171 (2000)
16. Abtin, F., Goldin, G.G., Brown, M., Rao, A., Kim, H., Jiang, A., Ahmad, S., Roback, D., Pais, R., Irani, Z.: Variation in Fiossure Anatomy in Emphysema Patients Treated with Endobronchial Valve. Posterboard presentation #33;A754 at ATS at Toronto (2008)