Published in final edited form as:
IEEE Trans Med Imaging. 2014 July ; 33(7): 1527–1540. doi:10.1109/TMI.2014.2317520.

Adaptive Quantification and Longitudinal Analysis of Pulmonary Emphysema with a Hidden Markov Measure Field Model

Yrjö Häme[*],
Columbia University, Department of Biomedical Engineering, New York, NY, USA

Elsa D. Angelini,
Telecom ParisTech, Institut Mines-Telecom, LTCI CNRS, Paris, France and with Columbia University, Department of Biomedical Engineering, New York, NY, USA

Eric A. Hoffman,
University of Iowa, Department of Radiology, Iowa City, IA, USA

R. Graham Barr, and
Columbia University, College of Physicians and Surgeons, Department of Medicine, New York, NY, USA

Andrew F. Laine
Columbia University, Department of Biomedical Engineering, New York, NY, USA

Abstract

The extent of pulmonary emphysema is commonly estimated from CT images by computing the proportional area of voxels below a predefined attenuation threshold. However, the reliability of this approach is limited by several factors that affect the CT intensity distributions in the lung.

This work presents a novel method for emphysema quantification, based on parametric modeling of intensity distributions in the lung and a hidden Markov measure field model to segment emphysematous regions. The framework adapts to the characteristics of an image to ensure a robust quantification of emphysema under varying CT imaging protocols and differences in parenchymal intensity distributions due to factors such as inspiration level. Compared to standard approaches, the present model involves a larger number of parameters, most of which can be estimated from data, to handle the variability encountered in lung CT scans.

The method was used to quantify emphysema on a cohort of 87 subjects, with repeated CT scans acquired over a time period of 8 years using different imaging protocols. The scans were acquired approximately annually, and the data set included a total of 365 scans. The results show that the emphysema estimates produced by the proposed method have very high intra-subject correlation values. By reducing sensitivity to changes in imaging protocol, the method provides a more robust estimate than standard approaches. In addition, the generated emphysema delineations promise great advantages for regional analysis of emphysema extent and progression, possibly advancing disease subtyping.

[*](yh2475@columbia.edu).

Keywords

Computed tomography; lung; image segmentation; Markov random fields; emphysema index

I. Introduction

Emphysema is a condition involving alveolar wall destruction [1]. A mixture of emphysema and small airways disease contributes to chronic airflow limitation, characteristic of chronic obstructive pulmonary disease (COPD), which is a leading cause of morbidity and mortality worldwide [2].

Computed tomography (CT) is commonly used to identify and quantify the extent of pulmonary emphysema. Generally, an estimate of emphysema severity is obtained using a densitometric measure, called *percent emphysema* (%emph) (also referred to as *emphysema index* or *percent low attenuation area*), which quantifies the proportion of voxels with intensity values below a fixed threshold within the lung region. The %*emph* measure was originally derived from the *density mask* [3], and is commonly used in clinical studies, but there is no consensus on the intensity threshold value that should be used. The threshold values typically range from −950 to −910 Hounsfield Units (HU) (see review by Hoffman et al. [4]). Another commonly used measure, the percentile density (PD) quantifies a predefined percentile of the intensity distribution, and this measure has been found to be preferable in longitudinal studies [5], [6].

Standard measures are influenced by several factors that cause variations in the intensity distributions present in lung CT images, observed as different levels of noise, and variable intensity levels and distribution shapes. These factors include image reconstruction algorithm, slice thickness, scanner type and calibration, radiation dose, gravity and inspiration level [1].

Adaptive smoothing for normalization of image data prior to thresholding has been proposed as a solution for images with different noise levels [7]. The study showed promise in obtaining similar %*emph* values between low-dose and regular CT scans. This approach, however, still requires thresholding after the filtering operation, and may be susceptible to variations in intensity levels.

Recent studies have proposed solutions for the normalization of %*emph* measures to account for differences caused by changes in reconstruction algorithms and slice thickness [8], [9]. Correction of %*emph* based on lung volume has also been recommended [10], [11] to adjust for variations in inspiration level. These approaches consider only a part of the sources of variation, and since they correct the final %*emph* value, they do not provide voxel labels that can be useful when assessing the spatial distribution of emphysema.

Image texture analysis has been applied for supervised classification of emphysema [12], [13], [14]. These approaches require labeled data to train classifiers, and have not been shown to be robust to changes in imaging protocols.

Emphysema quantification methods that are robust to variations in image intensity distributions are required for two purposes: (1) analysis of large cohorts of patients from multiple databases for population-wide analysis of emphysema, and (2) longitudinal analysis of emphysema progression, which has been recognized as an area where more research is currently required [1].

We propose an approach for the quantification of emphysema that uses a Hidden Markov Measure Field (HMMF) model [15]. The HMMF model adds an intermediate continuous-valued labeling, called the *measure field*, to the standard Markov Random Field (MRF) models [16], [17]. MRF models have been used extensively for many problems in image analysis (see [18]), as they provide a convenient probabilistic framework for modeling local interactions of image pixels and for including prior spatial constraints to a segmentation process. In a preliminary version of this work [19], we applied the HMMF model to segment emphysematous regions in full-lung CT scans. The study showed that the HMMF model is robust to changes in CT image reconstruction algorithms. The application of the model requires parameterizing intensity distributions within the lung, and defining a spatial regularization weight, but manually labeled training data is not needed. The HMMF model has also been used in our previous work for liver tumor segmentation [20].

Refining and expanding from our preliminary study [19], this paper gives a fully detailed presentation of an improved HMMF model for emphysema segmentation, and demonstrates the performance of the model on a large longitudinal data set of lung CT scans. The data set includes repeated full-lung scans among participants in the Emphysema and Cancer Action Project (EMCAP) [21] who were later recruited into the Multi-ethnic Study of Atherosclerosis (MESA) COPD Study [22]. The full-lung scans in the two studies were acquired with different imaging protocols. We compared the performance of the presented method to the standard %*emph* at −950 HU (%$emph_{-950}$) and the 15th percentile density measure (denoted PD_{15}). In addition, %*emph* at −950 HU was evaluated with prior Gaussian filtering of images ($\%emph^{G}_{-950}$).

II. Methods and Data

A. Overview

The proposed HMMF model includes the following components intended to improve the quantification of emphysema over the standard %*emph* measure:

1. Likelihood functions are defined by modeling intensity distributions observed in the data. This approach accounts for the variability in intensity distribution shapes, caused by changes in imaging protocol, such as slice thickness, scanner type and calibration, radiation dose, and reconstruction algorithm.
2. The locations of the likelihood functions are allowed to vary, to account for patient- and scan-specific variations, due to differences in the inspiration level and average parenchymal density.
3. An image voxel is assumed to belong more likely to the same class as its neighboring voxels than to a different class. This assumption takes the image

structure into consideration, an aspect that is entirely ignored in standard histogram-based emphysema measures. The aim is to reduce classification errors due to overlapping likelihood functions, i.e. when there is a high level of uncertainty in the classification due to noise or poor contrast. Therefore, this improvement also accounts for changes in imaging protocol.

Given a CT volume, the lungs were segmented in the preprocessing step, as described in Section II-D2. Within the delineated lung region, a two-class HMMF model was used to automatically segment emphysematous regions from the healthy parenchyma. The segmentation was subsequently used to measure the extent of emphysema, by quantifying the proportional volume of emphysematous regions with respect to the entire lungs. The imaging protocol-dependent parameter values required by the model were learned from the CT image data, as explained in Section II-E. We begin by presenting the core of the segmentation method in the following sections.

B. Segmentation with the HMMF model

Let I denote the input image, Ω represents the image domain, and $r \in \Omega$ is an image voxel. The segmentation process involves two steps. The first step computes a continuous-valued Markov random vector field $q = [q_1, q_2]$, where each q_k corresponds to the value for class k. The second step then generates a binary label field f from q. The vector field q is constrained by $q_1(r) + q_2(r) = 1$, $q_1, q_2 \geq 0$, where $q_k(r)$ is the value at voxel r, for class k.

The vector field q represents an intermediate labeling and is assigned a prior distribution that enforces spatial regularity

$$P_q(q) = \frac{1}{K} exp\left[-\sum_C W_C(q)\right], \quad (1)$$

where C are spatial cliques of a selected neighborhood system, W_C are potential functions associated with C, and K is a positive normalizing constant. Here, 3D pairwise cliques in 26-connected neighborhoods were used.

The potential functions W_C are designed to measure the smoothness of q within the neighborhood defined by the clique $C = [r_1, r_2]$, at voxels r_1 and r_2. In this work, the potential $W_{r_1 r_2}$ between two values $q(r_1)$ and $q(r_2)$ was defined as (similarly to [20]):

$$W_{r_1 r_2}(q) = \frac{\lambda}{Z} exp\left[-\frac{d(r_1 r_2)^2}{2\sigma_W^2}\right] \sum_{k=1}^{2} (q_k(r_1) - q_k(r_2))^2, \quad (2)$$

where $d(r_1, r_2)$ is the Euclidean distance between r_1 and r_2, which takes voxel spacing into account, σ_W and λ are scalar constants, and Z is a normalization term that scales the exponentials to sum to one within the 26-connected neighborhood E_{rj} of any voxel r_j (the value is constant for a single image): $Z = \Sigma_{ri \in E_{rj}} exp\left[-d(r_i, r_j)^2 / 2\sigma_W^2\right]$. The parameter σ_W controls how fast the Markov weight decreases as a function of voxel distance. This parameter is important for 3D neighborhoods in anisotropic volumes where the slice thickness is large compared to the in-plane resolution.

The value of λ controls the weight of the Markovian prior with respect to the likelihood function (described below), and should be adapted to the image content and the targeted segmentation task. When image data is noisy, individual likelihood values are less reliable than for a less noisy case, and the model should be forced towards the prior by increasing the value of λ. In this study, the value of λ was assigned based on the image reconstruction algorithm, as explained in Section II-E4.

For the segmentation process, the intensity distributions in an image I were modeled with parametric distributions $v_{\theta k}$ (detailed in Section II-C) where θ_k is the mean of the distribution for class k. The values θ_k have to be estimated simultaneously with q, and are assigned a prior distribution $P_\theta(\theta)$, defined in Section II-E1.

For a given image I, the posterior distribution for q and the associated parameter vector $\theta = [\theta_1, \theta_2]$ is obtained from the Bayes rule:

$$P(q,\theta|I) = \frac{1}{R} P(I|q,\theta)\, P_q(q)\, P_\theta(\theta), \quad (3)$$

where R is a positive normalizing constant.

The likelihood term in (3) is expressed as [15]:

$$P(I(r)|q,\theta) = v_{\theta_1}(r)\, q_1(r) + v_{\theta_2}(r)\, q_2(r). \quad (4)$$

Combining (1), (3), and (4), the respective *maximum a posteriori* (MAP) estimates q^* and θ^* for q and θ are found by maximizing $P(q,\theta|I) = \frac{1}{KR} exp[-U(q,\theta)]$, where

$$U(q,\theta) = -\sum_{r\in\Omega} log[v_{\theta_1}(r)\, q_1(r) + v_{\theta_2}(r)\, q_2(r)] + \sum_C W_C(q) - log(P_\theta(\theta)). \quad (5)$$

Since the normalization term KR is constant and positive, the MAP estimate is found by minimizing $U(q, \theta)$. The optimization was performed with the gradient projection Newtonian descent method, as formulated in [15].

Finally in the second step of the segmentation, a binary label field f^* was found by maximizing $P(f|q = q^*, \theta = q^*, I)$, which is simply done by finding the mode of each $q^*(r)$ (see [15]): $f^*(r) = 1$, if $q_1^*(r) > q_2^*(r)$, and $f^*(r) = 2$, otherwise. An example of the HMMF segmentation process is shown in Fig. 1.

C. Parametric functions for intensity distribution modeling

To obtain the likelihood values $P(I|q, \theta)$ in (4), we need to define the parametric functions $v_{\theta k}$ that model the intensity distributions for the two classes $k = 1, 2$. Ideally, these parametric distributions would have the same shape as the intensity distribution histograms for the two classes in the image data.

The intensity distributions within the lung were parameterized with skew-normal distributions:

$$v_{\theta_k} \equiv p\left(z_k, \alpha_k\right),$$

where α_k is a skew parameter. The auxiliary variable z_k is defined as:

$$z_k = \frac{I\left(r\right) - \theta_k}{\sigma_k},$$

where θ_k and σ_k represent the distribution location and scale, respectively.

The probability density function (pdf) for class k is defined as [23]:

$$p\left(z_k, \alpha_k\right) = 2\phi\left(z_k\right)\Phi\left(\alpha_k z_k\right), \quad (6)$$

where ϕ and Φ are the standard normal density and distribution function, respectively:

$$\phi\left(z\right) = \frac{1}{\sqrt{2\pi}} exp\left[-\frac{z^2}{2}\right]$$

and

$$\Phi\left(z\right) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right].$$

The notation erf refers to the error function:

$$erf\left(z\right) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

The values for the parameters α_k and σ_k were estimated from training data. To make the model adaptive, as explained in Section II-B, the parameter θ_k was allowed to vary for each individual image, while controlled by $P_\theta(\theta)$, as in (5).

D. Data and preprocessing

1) Database of CT scans—The evaluation data set consisted of 365 inspiratory chest CT scans from 87 subjects. The scans were collected in two different studies: in the EMCAP study [21], between 2004 – 2009, and subsequently in the MESA COPD study [22], between 2009–2011. The number of scans for each year is listed in Table I.

In addition to the evaluation data set, a parameter training set included 44 CT scans of 22 subjects from the EMCAP study. The 44 scans were acquired by reconstructing each of the 22 CT acquisitions with two different kernels, as detailed below. Three subjects in the

parameter training data set overlapped with the evaluation data set, but their scans were not used when tuning parameter values between imaging protocols.

The subjects in the data set were 50 – 79 years old, with 10 or more pack-year smoking history and who did not have clinical cardiovascular disease, stage IIIb-V kidney disease, asthma prior to age 45 years, other lung disease, prior lung resection, cancer, allergy to gadolinium, claustrophobia, metal in the body, pregnancy or weight > 300 lbs. [22]

In the EMCAP study, all subjects underwent non-contrast, full-lung CT scanning on a Siemens Sensation 16 scanner, with 120 kVp, a current between 169 mA and 253 mA, and speed 0.5 s. Of the total 278 scans in the evaluation data set, 259 were reconstructed with the B60f (sharp) convolution kernel, and 19 with the B31f (smooth) kernel. The 22 subjects in the parameter training set each had a single acquisition reconstructed with both the B31f and the B60f kernel, bringing the total number of training images to 44.

In the MESA COPD study, full-lung CTs were acquired with a GE LightSpeed VCT 64 scanner, at 120 kVp, 200 mA, 0.984 pitch, and speed of 0.5 s. Images were reconstructed with the standard (smooth) convolution kernel. This protocol is the same as the SPIROMICS/MESA Lung protocol [24], except that the mA was held fixed.

From all the available scans in the EMCAP study, we included all full-lung scans with a slice thickness of 0.75 mm. The EMCAP data in this study included 1 – 5 scans per patient, with at least 12 months between repeated scans. In the subsequent MESA COPD study, a single scan was acquired for each subject, bringing the total to 2 to 6 scans per subject.

The axial resolutions of the images used from the EMCAP data set were in the range [0.49, 0.87] mm and the slice thickness was 0.75 mm. For the MESA COPD data set, the axial resolution range was[0.58, 0.88] mm, and all scans had a slice thickness of 0.625 mm.

Fig. 2 illustrates the image appearance in the lung for the different imaging protocols, with detailed views of coronal slices from three scans of a single subject.

2) Preprocessing—Lungs and large airways were segmented from the background using an approach similar to [25], by applying an intensity threshold of −400 HU and then locating the largest connected objects in the resulting binary mask. Then, the trachea and some of the large airways were removed from the lung mask by closed space dilation [26].

The airway segmentation removed on average 0.9% (with standard deviation of 0.2%) of the initial mask volume. Since most of the volume of the airway segmentation corresponded to the trachea, we expect that any variability in the extent of the removed airways would have had only a minor effect on the resulting emphysema measure. In the experiments included in this paper, all the emphysema measures for a given scan were extracted using the same lung mask.

E. Model implementation and estimation of parameter values

1) Intensity distribution priors—For each class k, the locations θ_k of the parametric distributions $p(z_k, a_k)$ (6) are controlled by a prior distribution $P_\theta(\theta)$, which assigns

probabilities for different values of θ. The prior distribution affects the computation of the MAP estimate in (3) and (5). Using a uniform distribution for $P_\theta(\theta)$ means that the values of θ are driven entirely by the data, whereas a non-uniform distribution injects prior knowledge to their values, biasing the resulting MAP estimate θ^*.

In our segmentation task, the volume of emphysematous tissue in a given CT scan is unknown prior to the segmentation process. It is therefore important to ensure that θ_1, the distribution location of the emphysema class, does not receive unreasonably high values for healthier subjects due to a lack of samples in the emphysematous intensity range. On the other hand, there is always some healthy parenchyma present, and θ_2, which corresponds to the location of the distribution representing lung parenchyma and small vessels, can be allowed to vary more freely to fit the data.

The prior distribution $P_\theta(\theta)$ was assigned a delta function $\delta\left(\theta_1 - \mu^*_{\theta_1}\right)$ for θ_1, so that the value of θ_1 was fixed at $\mu^*_{\theta_1}$. The value was set as: $\mu^*_{\theta_1} = I_{air} + min\left(I^*_{tr} - I_{tr}, 0\right)$, where I_{air} = −1000 HU corresponds to the standard intensity of air outside the body in CT images, I_{tr} = −955 HU is a reference value of the tracheal air intensity, and I^*_{tr} is an imaging protocol-dependent tracheal air intensity estimate. I^*_{tr} was obtained by averaging over the airway segmentations generated in the preprocessing stage. For EMCAP, this was done with the parameter training set, and for MESA COPD using 20 randomly selected scans. This formula lowers the prior mean for class k = 1, if the intensity of tracheal air is lower than the calibration value. Wiemker et al. [27] have used tracheal air intensities similarly to adjust the intensity threshold for emphysema quantification. The values $\mu^*_{\theta_1}$, estimated for individual imaging protocols, are reported in Table II.

For θ_2, $P_\theta(\theta)$ was assigned a uniform distribution in the range [−995, −750] HU, to provide adaptivity. Due to the positive skewness ($\alpha_2 > 0$) of the distribution $p(z_2, \alpha_2)$ (see Section II-E2), the distribution peak is always located at a higher intensity than θ_2, and therefore the range of $P_\theta(\theta)$ cannot be directly interpreted as the range of possible mean intensities of the lung parenchyma.

2) Distribution parameter values—The parameterization of intensity distributions with the pdfs $p(z_k, \alpha_k)$, as defined in (6), involves defining values for σ_k and α_k for the two classes, with k = 1 and k = 2 representing the emphysematous tissue and healthy parenchyma, respectively.

To find the parameter values for class k = 2, subjects with the mildest cases of emphysema in the parameter training set were selected for each imaging protocol. For these subjects, the observed intensity values within the lung were assumed to correspond almost entirely to healthy lung parenchyma, with the highest intensity values caused by the partial volume effect from small vessels.

The scans representing mildest cases of emphysema within the parameter training set of each imaging protocol were selected as follows:

- EMCAP B31f: 10 subjects in the parameter training set with the lowest values of the standard %*emph* at −950 HU (%$emph_{-950}$) for B31f reconstructions, with all %$emph_{-950}$ < 5.0.
- EMCAP B60f: B60f reconstructions of the same 10 acquisitions as used for EMCAP B31f.
- MESA COPD: Randomly selected 10 subjects, which all had %$emph_{-950}$ < 1.0.

Skew-normal pdfs (6) were fitted to each of the normalized lung intensity histograms of the training scans within the range [−1000, −750] HU, as illustrated in Fig. 3. From the fitted skew-normal pdf, the estimated values α_2^* and σ_2^*, for the respective parameters α_2 and σ_2, were collected for each scan. For each imaging protocol, the average values of the parameter estimates over the training set (listed in Table II) were then used for the evaluation data set. The skew-normal pdfs with the mean α_2^* and σ_2^* are illustrated in Fig. 3(d) for the three imaging protocols.

The fit accuracy was measured by computing the histogram intersection d_H [28] between each intensity histogram *H(I)* and the estimated skew-normal pdf *v*. The histogram intersection is defined as: $d_H = 1 - \frac{1}{2}\Sigma_{j=1}|H(I)_j - v(I_j)|$, where *j* denotes histogram bins, and I_j is the center intensity of bin *j*. The range of d_H is from 0, corresponding to entirely non-overlapping histograms, to 1, for identical histograms.

For each imaging protocol, the average values of d_H over individual fits were (mean ± standard deviation): EMCAP B60f: 0.99±0.01; EMCAP B31f: ± 0.97±0.01; MESA COPD: 0.97 ± 0.02. With d_H evaluated between the histograms and the pdfs using the mean estimates α_2^* and σ_2^*, the average values over the training scans were: EMCAP B60f: 0.98±0.02; EMCAP B31f: 0.95 ± 0.02; MESA COPD: 0.92 ± 0.04. The histogram intersection values show that individual fits were very accurate, and the accuracy was fairly well retained when using the mean estimates α_2^* and σ_2^*. The variability in the MESA COPD intensity distribution shapes may be due to variable breath-hold levels or conditions that affect the density of the lungs.

The parameter values for class *k* = 1 (emphysema class) could not be estimated in the same manner as for class *k* = 2, since the intensity distributions of emphysematous voxels overlap with parenchymal intensity distributions. Also, since CT image intensities were limited to be higher than −1024 HU, finding a proper parameterization of the distribution from data could prove challenging even if reliable delineations were available.

Since the appropriate values of σ_k are affected mostly by image noise, it seems reasonable to assume that $\sigma_1 \sim \sigma_2$. The exact correspondence was not investigated in this study, and the parameter was assigned as $\sigma_1 = \sigma_2$. The shape of the class *k* = 1 intensity distribution is also unknown, and the skew parameter was set to $\alpha_1 = 0$, making the parametric distribution $p(z_1, \alpha_1)$ a standard normal distribution.

3) Model initialization—To minimize computational cost and to simplify the optimization process, the initial values $\theta^0 = \left[\theta_1^0, \theta_2^0\right]$ for θ were chosen with the aim that they would be close to their final values, on average, for each imaging protocol. For class $k = 1$, the value was assigned simply as the location of the prior distribution: $\theta_1^0 = \mu_{\theta_1}^*$. The values θ_2^0 were determined by fitting a skew-normal distribution in the range [−1000, −750] HU on a training data set, using the estimated parameters σ_2^* and α_2^* (see Table II), and taking the median value of the resulting locations. For EMCAP, this estimation was performed using the 22 images in the parameter training data set, and for MESA COPD, the same 20 randomly selected scans as in II-E1 were used. The resulting values of θ^0 are reported in Table II.

The initial values $q^0 = \left[q_1^0, q_2^0\right]$ for q were assigned at each voxel r using the initial values θ^0:

$$q_1^0(r) = \frac{v_{\theta_1^0}(r)}{v_{\theta_1^0}(r) + v_{\theta_2^0}(r)},$$

and $q_2^0(r) = 1 - q_1^0(r)$.

4) Markov field regularization parameters—The value of the Markov field weight λ should be assigned based on the level of noise in the image. The value of λ was tuned between the EMCAP B31f and B60f scans. For MESA COPD scans, λ was assigned to be same as for EMCAP B31f, due to similar noise levels. Of the 22 subjects in the EMCAP parameter training set, the 3 subjects included in the evaluation data set were removed when optimizing the λ parameter, to separate training and evaluation sets.

First, the CT scans reconstructed with the smooth kernel (B31f) were segmented with a low value for λ, namely λ = 1.0. This provided reference emphysema measures, denoted $\%emph_{MF}^{31}$. Then, for the sharp reconstructions (B60f) of the same CT acquisitions, λ was varied and the absolute differences between the resulting $\%emph_{MF}^{60}$ and the corresponding $\%emph_{MF}^{31}$ were computed. Finally, the B60f reconstructions were assigned a value for λ that minimized the mean absolute difference (MAD) over the training data set.

The MAD values and their standard deviations for different values of θ are shown in Fig. 4. The minimum MAD was 0.5, with a standard deviation of 0.4. This value was obtained using λ = 5.0 for the B60f images, and the other parameter values as presented in Table II. Fig. 5 illustrates the effects of varying the value of λ, on two image reconstructions.

Related to the λ parameter, the parameter σ_W (2) controls the decrease of the Markov weight with respect to distance. Since the slice thicknesses in the scans used in this study were similar to the in-plane resolutions, σ_W had only a minor influence on the results. Nonetheless, it has the desired effect of reducing the weight at the corners of the neighborhood. The same value σ_W = 1.5 mm as in [20] was used.

F. Instructions for model employment

To conclude the Section, we provide step-by-step instructions to employ the presented model on a new full-lung CT data set. The steps should be applied sequentially and separately for each imaging protocol in the data set.

1. Apply preprocessing on all the CT scans to generate segmentations of the lungs and the main airways.
2. Select scans with the least emphysema (for example by choosing the 10 subjects with the lowest $\%emph_{-950}$). Fit a parametric distribution to each of the histograms of the selected scans and collect the parameter values to model the likelihood function for the parenchyma class. In this study, skew-normal distributions were used, so this step produced estimates for the scale (σ_2^*) and skew (α_2^*) parameters. The distributions learned from mild cases in this step are expected to provide accurate likelihood values for the parenchyma class.
3. Select scans randomly from the data set. Fit the parametric distribution obtained in the previous step on the intensity histograms and find an estimate of the median of the parenchymal distribution location θ_2, and use this as the initialization value θ_2^0. Choosing an initialization value near the data set average provides more reliable results and faster optimization than using a predefined initial value.
4. Estimate the mean tracheal intensities using the airway segmentations of the scans in the previous step. Use the tracheal intensity mean to define the intensity distribution location θ_1 for the emphysema class. Airway intensity values provide an indication of average emphysema intensity values, and this information is used to provide accurate likelihood values for the emphysema class.
5. Assign a value for the Markov field weight λ. If multiple reconstructions of CT acquisitions are available, the approach used in the present study can be replicated. This requires defining a low λ value for the smoothest scans and adjusting the value for noisier scans, by minimizing the mean absolute difference of $\%emph_{\mathrm{MF}}$ between the reconstructions. Alternatively, the weight value can be inferred from results in previous studies, based on the parametric distribution scale or some measure of image noise, such as the local noise estimation in [7]. The Markov weight is intended to improve the segmentation results by reducing uncertainty caused by image noise. Here, the Markov field is implemented with a 3D neighborhood to enforce segmentation regularity between image slices, as well as within them. For scans with thicker slices, 2D neighborhoods may suffice (i.e. not enforcing regularity across slices).
6. Finally, using the parameter values learned in the previous steps, initialize the HMMF model and apply the segmentation method. The values of $\%emph_{\mathrm{MF}}$ are obtained by computing the volume classified as emphysema, divided by the total lung volume.

III. Results

A. Average emphysema scores over the evaluation database

In the following, %*emph*$_{\mathrm{MF}}$ refers to the %emph measure obtained with the presented HMMF model. %*emph*$_{-950}$ represents the standard %*emph* using a threshold of −950 HU, which is commonly used in clinical studies [29], [30], and PD_{15} is the 15^{th} percentile density.

The %*emph* measure obtained by thresholding at −950 HU after 3D Gaussian filtering is denoted as $\%emph^{G}_{-950}$. The filter scale σ_G was optimized in the same way as the value of λ, by minimizing the MAD on the parameter training data set (see Section II-E4 for details). For Gaussian filtering, MAD was minimized at 1.1 ± 1.3, using σ_G = 0.74 (see Fig. 4).

The values provided by these four emphysema measures are generally referred to as *emphysema scores*. All %*emph* scores are reported in the range [0, 100], corresponding to percentages of total lung volume, and PD_{15} scores are reported in HU.

A general overview of the emphysema scores over the evaluation data set of 87 subjects is provided in Table III, with mean values, standard deviations, and minimum and maximum values reported using the four emphysema measures, for each imaging protocol. In addition, the values are reported using the most recently acquired EMCAP B60f scan for each subject. For 49 subjects, the most recent B60f scan was acquired in 2008-09, for 35 subjects in 2007 and for the remaining 3 subjects in 2006. This scan grouping enables a comparison to the same population of 87 subjects that was available in MESA COPD.

Based on the assumption that emphysema is irreversible, %*emph* should theoretically not decrease with time. Since the majority of the population in this study represented mild cases of emphysema, only a minor increase in the mean of %*emph* values was expected. The PD_{15} measure should decrease slightly for the same reasons.

The overall statistics show that the average %*emph*$_{\mathrm{MF}}$ remained fairly stable, while the average %*emph*$_{-950}$ varied greatly depending on the imaging protocol. Indeed, between imaging protocols the mean of %*emph*$_{\mathrm{MF}}$ increased by 1.7 from the most recent EMCAP B60f scans to the MESA COPD scans, while the mean of %*emph*$_{-950}$ decreased by 29.6 for the same data. With prior Gaussian smoothing ($\%emph^{G}_{-950}$) on B60f scans, the mean decreased by 3.4. The PD_{15} measure increased by 95 HU. In a paired t-test, all these changes were statistically significant at the 5% level.

As the EMCAP B31f and MESA COPD scans were reconstructed with smooth kernels and acquired 1 or 2 years apart, they were expected to give similar emphysema scores. However, the mean values of %*emph*$_{-950}$ declined significantly, from 8.5 to 2.6, and the mean PD_{15} values increased by 22 HU. In comparison, %*emph*$_{\mathrm{MF}}$ showed only a slight increase for the same data set, from 4.7 to 5.5. Also these changes were statistically significant at the 5% level.

B. Pairwise intra-measure correlations between longitudinal scans

Pairwise correlations between emphysema scores from longitudinal scans of individual subjects were computed. The results are reported in Table IV (evaluations with fewer than 16 cases were omitted for space considerations). In the following, comparisons of correlations were performed using Fisher's r-to-z transformation and a two-tailed test of the resulting z-score. Fig. 6 shows scatterplots of the emphysema scores from EMCAP 2008 – 09 and MESA COPD.

The results show that %$emph_{\mathrm{MF}}$ achieved very high correlations between longitudinal scans regardless of the imaging protocol. All 10 comparisons with more than 20 scans had correlations of 0.95 or higher, whereas the overall minimum was 0.85. In only one comparison (18 scans between 2004 and 2007, B60f reconstruction) the correlation for %$emph_{\mathrm{MF}}$ was lower than for another measure (PD_{15}), but the difference was not statistically significant ($p = 0.92$).

While the correlation values for %$emph_{-950}$ were relatively high when comparing scores between EMCAP B60f scans, in the range [0.73, 0.89], their values declined significantly when comparing scores from B60f scans to scores from EMCAP B31f scans [0.64, 0.69], or to scores from MESA COPD scans [0.52, 0.70]. Gaussian filtering of B60f scans before thresholding improved the correlation values, particularly for comparisons to MESA COPD scans. Interestingly, correlations between B60f scans were also higher for $\%emph^{G}_{-950}$ than for %$emph_{-950}$. Correlations of PD_{15} were similar to or lower than correlations of %$emph_{-950}$.

When comparing EMCAP B60f scores to MESA COPD scores, all correlation values were higher for %$emph_{\mathrm{MF}}$ than for any other measure. This difference was significant ($p < 0.01$) for all comparisons, except for EMCAP B60f 2004 ($p = 0.07$), where the sample size was the smallest ($N = 23$). Between EMCAP B31f and MESA COPD, %$emph_{\mathrm{MF}}$ also had the highest correlation (0.96), but the difference to the PD_{15} value (0.87) was not statistically significant ($p = 0.08$) due to the small sample size ($N = 19$).

C. Correlations between emphysema measures

To study the correspondence between %$emph_{\mathrm{MF}}$ and %$emph_{-950}$, pairwise correlations were computed. The results are reported in Table V.

The correlation values show that there was a good agreement between %$emph_{\mathrm{MF}}$ and %$emph_{-950}$ for the MESA COPD scans. For the EMCAP scans, the correlation values were significantly lower. When taking into account the high intra-subject correlations of %$emph_{\mathrm{MF}}$ in Table IV, the high correlation for the MESA COPD data indicates that %$emph_{\mathrm{MF}}$ values from the EMCAP scans are also comparable to the MESA COPD %$emph_{-950}$ values.

D. Progression of emphysema measures

Subject-specific differential %*emph* scores (δ%*emph*(t_E)) were generated by subtracting the MESA COPD %*emph* value (%*emph*(t_{MC})) from all preceding (EMCAP) %*emph* values (%*emph*(t_E)) of the same subject:

$$\delta\%emph\,(t_E) = \%emph\,(t_E) - \%emph\,(t_{MC})\,.$$

This way, negative values of δ%*emph*(t_E) indicate growth over time.

The mean values and standard deviations of δ%*emph*(t_E) for the three %*emph* measures are shown in Fig. 7. The annual progression rate of %$emph_{-950}$ has been previously estimated as 0.63 (SE 0.03) [31], albeit for a different patient population. This estimate is used as a reference progression rate in Fig. 7.

The figure shows that on average, %$emph_{\mathrm{MF}}$ increased steadily and the differential scores had a relatively low standard deviation throughout the studied data set. Also, the average progression rate of %$emph_{\mathrm{MF}}$ seems to agree well with the progression rate reported in [31]. For the standard measure %$emph_{-950}$, the values of δ%$emph_{-950}$(t_E) were large and had high variability. With prior Gaussian filtering of B60f scans, $\%emph^{G}_{-950}$ had better agreement with the MESA COPD values than %$emph_{-950}$, but the differential scores still had high variability. Moreover, the %$emph_{-950}$ values from MESA COPD were clearly lower than $\%emph^{G}_{-950}$ values from EMCAP, suggesting a decrease of %*emph* in time (see also Table III).

To evaluate annual changes Δ_a%*emph* in the emphysema measures, the changes in %*emph* between consecutive scans were computed and divided by the number of years between the scans at time points t_1 and t_2 (equivalently for $\Delta_a PD_{15}$): Δ_a%*emph* = (%*emph*(t_2) − %*emph*(t_1))/($t_2 - t_1$). Histograms of the annual changes are shown in Fig. 8.

The histograms show that the values of Δ_a%$emph_{\mathrm{MF}}$ were centered close to 0, with more instances in the positive values, indicating an increase in %$emph_{\mathrm{MF}}$. Out of 278 evaluations of Δ_a%$emph_{\mathrm{MF}}$, 81% were in the range [−1, 2]. On the other hand, the values of Δ_a%$emph_{-950}$ and $\Delta_a PD_{15}$ had very wide and irregular distributions.

The annual progression rates were computed by taking the mean of Δ_a%*emph* for each subject. First, the annual progression rate was evaluated using only the EMCAP B60f data set. The means and standard deviations of the annual progression rates over the 87 subjects for the four emphysema measures in EMCAP B60f were: %$emph_{\mathrm{MF}}$: 0.54 ± 0.91, %$emph_{-950}$: 2.15 ± 2.60, $\%emph^{G}_{-950}$: 0.51 ± 2.67, and PD_{15} : −16.3 ± 13.1. The results show very similar progression rates between %$emph_{\mathrm{MF}}$ and $\%emph^{G}_{-950}$, but the latter suffers from higher variability.

The overall annual progression rate of %$emph_{\mathrm{MF}}$ was computed over the entire evaluation data set. The mean and standard deviation were 0.56 ± 1.37. The progression rate depended on the degree of emphysema. Out of the 87 subjects, 32 had %$emph_{\mathrm{MF}}$ above 5.0 for the

MESA COPD scan, and for these subjects the mean and standard deviation of $\Delta_a\%emph_{\mathrm{MF}}$ were 1.3 ± 1.0. For the remaining 55 milder cases, the corresponding values were 0.4 ± 0.3.

For reference, a study by Parr et al. [5] estimated the annual progression rate of $\%emph_{-950}$ within a range [0.1, 2.0], and the annual progression rate of PD_{15} within [−1.2, −1.9]. The progression rate depended on disease status, so that more severe disease progressed faster. The presented results for the annual progression rate of $\%emph_{\mathrm{MF}}$ agree with these estimates, as well as with the reference value in Fig. 7.

E. Example case

Fig. 9 presents an example of emphysema masks and the associated emphysema scores from six scans for a single subject. The figure illustrates how $\%emph_{\mathrm{MF}}$ increased gradually with time, while $\%emph_{-950}$ depended largely on the image acquisition protocol. The figure also exemplifies the quality of the emphysema masks produced by the HMMF segmentation, as the emphysema regions are seemingly consistent across the different scans.

F. Sensitivity to parameter value estimation

The sensitivity of the HMMF model with respect to the σ_2, α_2 and θ_2^0 parameter values was evaluated with 22 B60f and B31f scans in the EMCAP training set, and a randomly selected subset of 20 scans from the MESA COPD data set, where mean ± standard deviation of $\%emph_{\mathrm{MF}}$ were 5.4±5.2. The changes in $\%emph_{MF}$ were quantified when adjusting each parameter value, while keeping other values unchanged. The changes $\Delta\sigma_2$ for σ_2 and $\Delta\alpha_2$ for α_2 were given values of ±1, ±2 times the standard deviations of the respective estimates σ_2^* and α_2^* for each imaging protocol, as listed in Table II. Since θ_2^0 was estimated by taking the median over randomly selected scans, the range of values $\Delta\theta_2^0$ was defined by estimating θ_2^0 25 times (see Section II-E3) on EMCAP B60f and on MESA COPD, each time with a different randomly selected set of 20 scans (for EMCAP B31f, not enough scans were available for repeating the estimation). The standard deviation of the resulting θ_2^0 estimates was 4.1 HU for EMCAP B60f and 4.5 HU for MESA COPD (the respective means were −981 HU and −933 HU). Therefore, $\Delta\theta_2^0$ was assigned values between −10 and 10 HU, to approximate the range of the first two multiples of the estimate standard deviation.

The resulting changes in $\%emph_{\mathrm{MF}}$ are shown in Fig. 10. The model was more sensitive with respect to θ_2^0 than the other two parameters. When the change in σ_2 or α_2 was within one standard deviation of the parameter value estimate, the absolute mean change in $\%emph_{\mathrm{MF}}$ was less than 1.0, for all imaging protocols. Adjusting θ_2^0 by ±5 HU also changed the mean $\%emph_{\mathrm{MF}}$ by less than 1.0, except for EMCAP B60f, where increasing θ_2^0 by 5 HU resulted in a mean change of 1.1.

Of the three data sets, MESA COPD showed the most sensitivity with respect to σ_2, although this is at least partially caused by the high standard deviation associated with the σ_2 estimates. EMCAP B31f scans displayed the least sensitivity with respect to θ_2^0, while the most sensitive were EMCAP B60f scans. The sensitivity of the EMCAP B60f scans is assumed to be due to the high Markov field weight λ used for this imaging protocol. With a

high value of λ, spatial regularity is strongly enforced, causing the initialization to have more influence on the result than with a lower value of λ.

These results show that the presented model is somewhat sensitive to the parenchyma class location initialization (θ_2^0), although this value can be estimated fairly consistently for a given imaging protocol. Decreasing the value of θ_2^0 resulted in a smaller absolute change in %*emph*$_{\mathrm{MF}}$ than increasing its value. The resulting %*emph*$_{\mathrm{MF}}$ always increased with an increase in θ_2^0, indicating that changes in θ_2^0 alter %*emph*$_{\mathrm{MF}}$ in a consistent direction. The value of θ_2^0 could therefore be used to adjust the sensitivity of %*emph*$_{\mathrm{MF}}$, and provide lower and upper bounds for the extent of emphysema, as proposed for tumoral growth estimates in [32].

G. Computational expense

The computational expense of the HMMF model depends on the number of voxels within the lung segmentation and the convergence speed of the optimization process. A typical lung region in this study included between $10 \cdot 10^6$ and $25 \cdot 10^6$ voxels. With the current C programming language implementation, the HMMF segmentation was computed for 100 images in 407 minutes using four computing cores, corresponding to an average computation time of 16.3 min/N per image, where N is the number of cores used. We expect code optimization to reduce the required computation time.

IV. Discussion and future work

This study presented a novel method for the quantification of emphysema from lung CT images. The method is based on a segmentation of emphysematous regions from the lung parenchyma with a HMMF model. This approach is analogous to the original *density mask* [3], which has become popular in clinical studies. The presented segmentation model introduces a parameterization of the intensity distributions and a probabilistic labeling of voxels that enforces spatial coherence of the resulting label regions. These qualities were shown to provide segmentations that were robust to changes in imaging protocols, and subsequently enabled consistent and robust quantification of emphysema with the %*emph*$_{\mathrm{MF}}$ measure.

The presented method was shown to be valuable for quantifying emphysema in a longitudinal data set where imaging protocols and CT scanners changed over time. Using CT scans from the EMCAP and the MESA COPD studies, the results showed that thresholding-based %*emph*$_{-950}$ values were not comparable between the two studies, whereas %*emph*$_{\mathrm{MF}}$ values showed good agreement. Prior Gaussian filtering improved the thresholding-based measure on noisy scans, but the correlation values were still lower than for %*emph*$_{\mathrm{MF}}$. Interestingly, %*emph*$_{\mathrm{MF}}$ also resulted in higher intra-subject correlations than %*emph*$_{-950}$ for longitudinal scans acquired with a single imaging protocol.

For the MESA COPD scans, the %*emph*$_{\mathrm{MF}}$ values were on average higher than the values for %*emph*$_{-950}$. This means that %*emph*$_{\mathrm{MF}}$ would correspond on average to a standard thresholding-based %*emph* measure using a higher threshold value than −950 HU. However,

this does not mean that a higher threshold would replicate the HMMF masks, or that the $\%emph_{MF}$ values could be replicated by simply identifying an intensity threshold value for each imaging protocol. The presented HMMF model adapts to the intensity distribution of each individual scan to provide a unique separation of the emphysematous regions from the lung parenchyma, while imposing a spatial regularization that goes beyond labeling by pure intensity-based analysis.

Even though the values of $\%emph_{MF}$ were higher than $\%emph_{-950}$ on average for the MESA COPD data, the correlation between the two measures was high, indicating a good agreement for this imaging protocol. The average %*emph* values from MESA COPD scans for the studied population were relatively low. While this study showed that the HMMF segmentations are able to provide consistent %*emph* values between imaging protocols, a future study should be performed on a population with more severe cases of emphysema.

The annual progression rate of $\%emph_{MF}$ in this study was similar to what has been found previously for $\%emph_{-950}$ in other studies [5], [31]. In comparison, the estimated annual progression rate of $\%emph_{-950}$ in the present study was significantly affected by changes in imaging protocols, and did therefore not correspond to the estimates found in studies using data acquired with a single imaging protocol. Also, for $\%emph_{MF}$ the progression rate in mild cases of emphysema was found to be slower than for severe cases.

For the EMCAP scans, the mean of $\%emph_{-950}$ showed a relatively large average annual increase. However, this change might not be indicative of emphysema progression, as the mean of $\%emph_{-950}$ for the latest scans in MESA COPD was relatively low. Still, the intra-subject correlations for $\%emph_{-950}$ in the EMCAP data remained relatively high. These qualities suggest that the $\%emph_{-950}$ values for the EMCAP scans hold patient-specific information, but their absolute values should be used with caution, and the differences between longitudinal scans may be mostly due to changes in image acquisition protocols.

Generating emphysema masks using a robust and consistent segmentation method may have significant value beyond merely extracting a single estimate of emphysema extent. The masks provide information needed to assess the spatial distribution and regional progression of emphysema. The presented model enforces smoothness of the emphysema masks, which is particularly important for scans with high levels of noise. This may prove very valuable for morphological analysis of segmentation masks, used for emphysema quantification and subtyping [33], [34], [35]. Visually, the generated HMMF-based emphysema masks seemed to correspond to each other between longitudinal scans. Our future work will include intra-subject registration of scans to quantify the overlap of the generated emphysema masks, and their regional evolution on longitudinal data.

One of the shortcomings of the HMMF model is the requirement to learn parameter values for each CT imaging protocol. Estimation of the parenchymal intensity distribution parameters requires either normal subjects or mild cases of emphysema. In our preliminary work [19], the intensity distributions were modeled as normal distributions. However, normal distributions were not able to account for the typical heavy tails of the intensity distributions towards higher parenchymal intensities. This led to poor fits that often affected

the final location of the parametric distribution, and caused unreliable estimates of emphysema. By adding the skew parameter that is fairly simple to estimate, the skew-normal distribution improves the fit to the data. The method was shown to be somewhat sensitive to the initial value of the parenchymal likelihood function location, but with a predictable effect on the resulting emphysema estimate. This issue will be investigated in future work, by studying the effect of the initial value on the intermediate measure field. Also, sensitivity might be reduced by adopting a re-initialization scheme after the initial optimization of the distribution location for a given scan. Another drawback of the method is the computational cost of generating the HMMF segmentation, which is obviously higher than for simple thresholding. However, with currently available computational resources this should not be a critical issue, even though real-time processing does not seem achievable.

Further development is still possible to improve the segmentation of emphysematous regions from lung CT scans. In particular, gravity often causes the average intensity within the lung parenchyma to vary spatially. Sometimes this unevenness can affect the thresholding-based %*emph* values as the intensity values may decrease below the set threshold and cause an over-estimation of emphysema. The presented version of the HMMF model does not fully alleviate this problem, as the intensity distribution is modeled globally over the whole lung region. Future development will focus on regional intensity distribution modeling, to adjust the lung parenchyma intensity mean estimate according to the effect of gravity. Alternatively, the CT intensity values could be adjusted for the effect of gravity in the preprocessing stage, as proposed in [36]. Future work will also include a comparison study with other emphysema quantification approaches [7], [8], [9].

The lack of ground truth is a shortcoming of the evaluation performed in this study. Establishing ground truth would require repeated pathological sections, which is not feasible in humans. However, this study showed that the proposed method can be used to obtain robust and replicable estimates of emphysema extent across imaging protocols, which is a prerequisite for further study of their clinical relevance.

While providing extraordinary data for diagnostic purposes, the increase in repeated CT scans for patient monitoring has raised concerns about imaging-based health risks caused by radiation. The presented emphysema quantification method may prove valuable for accurately quantifying emphysema even as image noise levels are elevated when reducing scanner radiation doses. The method was already shown to improve the quantification of emphysema on existing heterogeneous image data, enabling better understanding of the disease.

Acknowledgments

The authors thank Research Professor Tuomas Häme of VTT Technical Research Center of Finland for technical advice and valuable comments. The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org. The authors also thank the anonymous reviewers for their valuable comments.

Manuscript was submitted for review to IEEE Transactions on Medical Imaging on May 31st, 2013. Y. Häme is funded by the International Fulbright Science and Technology Award at the U.S. Department of State. CT scans were acquired with funding from NIH/NHLBI R01-HL075476 and R01-HL093081 with additional support from R01-HL077612.

References

[1]. Mets OM, de Jong PA, van Ginneken B, Gietema HA, Lammers JWJ. Quantitative computed tomography in COPD: possibilities and limitations. Lung. 2012:1–13.

[2]. Global initiative for chronic obstructive lung disease. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. Updated 2011, accessed October 2012. Available at: http://www.goldcopd.com

[3]. Müller NL, Staples CA, Miller RR, Abboud RT. Density mask. An objective method to quantitate emphysema using computed tomography. CHEST Journal. 1988; 94(4):782–787.

[4]. Hoffman EA, Simon BA, McLennan G. State of the art. A structural and functional assessment of the lung via multidetector-row computed tomography phenotyping chronic obstructive pulmonary disease. Proceedings of the American Thoracic Society. 2006; 3(6):519–532. [PubMed: 16921136]

[5]. Parr DG, Stoel BC, Stolk J, Stockley RA. Validation of computed tomographic lung densitometry for monitoring emphysema in 1-antitrypsin deficiency. Thorax. 2006; 61(6):485–490. [PubMed: 16537666]

[6]. Newell JD, Hogg JC, Snider GL. Report of a workshop: quantitative computed tomography scanning in longitudinal studies of emphysema. European Respiratory Journal. 2004; 23(5):769–775. [PubMed: 15176695]

[7]. Schilham AMR, van Ginneken B, Gietema H, Prokop M. Local noise weighted filtering for emphysema scoring of low-dose CT images. IEEE Transactions on Medical Imaging. 2006; 25(4):451–463. [PubMed: 16608060]

[8]. Ceresa M, Bastarrika G, de Torres JP, Montuenga LM, Zulueta JJ, Ortiz-de Solorzano C, Munoz-Barrutia A. Robust, standardized quantification of pulmonary emphysema in low dose CT exams. Academic Radiology. 2011; 18(11):1382–1390. [PubMed: 21852160]

[9]. Bartel ST, Bierhals AJ, Pilgram TK, Hong C, Schechtman KB, Conradi SH, Gierada DS. Equating quantitative emphysema measurements on different CT image reconstructions. Medical Physics. 2011; 38(8):4894–4902. [PubMed: 21928661]

[10]. Stoel BC, Putter H, Bakker ME, Dirksen A, Stockley RA, Piitulainen E, Russi EW, Parr D, Shaker SB, Reiber JHC, Stolk J. Volume correction in computed tomography densitometry for follow-up studies on pulmonary emphysema. Proceedings of the American Thoracic Society. 2008; 5(9):919–924. [PubMed: 19056717]

[11]. Shaker SB, Dirksen A, Laursen LC, Skovgaard LT, Holstein-Rathlou N-H. Volume adjustment of lung density by computed tomography scans in patients with emphysema. Acta Radiologica. 2004; 45(4):417–423. [PubMed: 15323394]

[12]. Sorensen L, Nielsen M, Lo P, Ashraf H, Pedersen JH, de Bruijne M. Texture-based analysis of COPD: A data-driven approach. IEEE Transactions on Medical Imaging. 2012; 31(1):70–78. [PubMed: 21859615]

[13]. Ginsburg SB, Lynch DA, Bowler RP, Schroeder JD. Automated texture-based quantification of centrilobular nodularity and centrilobular emphysema in chest CT images. Academic Radiology. 2012; 19(10):1241–1251. [PubMed: 22958719]

[14]. Xu Y, Sonka M, McLennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. IEEE Transactions on Medical Imaging. 2006; 25(4):464–475. [PubMed: 16608061]

[15]. Marroquin JL, Santana EA, Botello S. Hidden Markov measure field models for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003; 25(11): 1380–1387.

[16]. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984; 6(6):721–741. [PubMed: 22499653]

[17]. Besag J. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society. Series B (Methodological). 1986:259–302.

[18]. Li, SZ. Markov random field modeling in image analysis. Springer; 2009.

[19]. Häme, Y.; Angelini, ED.; Hoffman, EA.; Barr, RG.; Laine, AF. Robust quantification of pulmonary emphysema with a hidden Markov measure field model; IEEE 10th International Symposium on Biomedical Imaging (ISBI); 2013; p. 382-385.

[20]. Häme Y, Pollari M. Semi-automatic liver tumor segmentation with hidden Markov measure field model and non-parametric distribution estimation. Medical Image Analysis. 2012; 16(1):140–149. [PubMed: 21742543]

[21]. Mesia-Vela S, Yeh C-C, Austin JHM, Dounel M, Powell CA, Reeves A, Santella RM, Stevenson L, Yankelevitz D, Barr RG. Plasma carbonyls do not correlate with lung function or computed tomography measures of lung density in older smokers. Biomarkers. 2008; 13(4):422–434. [PubMed: 18484356]

[22]. Thomashow MA, Shimbo D, Parikh MA, Hoffman EA, Vogel-Claussen J, Hueper K, Fu J, Liu C-Y, Bluemke DA, Ventetuolo CE, Doyle MF, Barr RG. Endothelial microparticles in mild chronic obstructive pulmonary disease and emphysema. The Multi-Ethnic Study of Atherosclerosis Chronic Obstructive Pulmonary Disease study. American Journal of Respiratory and Critical Care Medicine. 2013; 188(1):60–68. [PubMed: 23600492]

[23]. Azzalini A. A class of distributions which includes the normal ones. Scandinavian journal of statistics. 1985:171–178.

[24]. Sieren JP, Hoffman EA, Baumhauer H, Barr RG, Goldin JG, Rennard S. CT imaging protocol standardization for use in a multicenter study: SPIROMICS. Radiological Society of North America. 2011

[25]. Hu S, Hoffman EA, Reinhardt JM. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. IEEE Transactions on Medical Imaging. 2001; 20(6):490–498. [PubMed: 11437109]

[26]. Masutani, Y.; Masamune, K.; Dohi, T. Visualization in Biomedical Computing. Springer; 1996. Region-growing based feature extraction algorithm for tree-like objects; p. 159-171.

[27]. Wiemker, R.; Bülow, T.; Blaffert, T.; Dharaiya, E. SPIE Medical Imaging. Vol. 7263. International Society for Optics and Photonics; 2009. Correlation of emphysema score with perceived malignancy of pulmonary nodules: A multi-observer study using the LIDC-IDRI CT lung database.

[28]. Cha S-H. Comprehensive survey on distance/similarity measures between probability density functions. City. 2007; 1(2):1.

[29]. Galbán CJ, Han MK, Boes JL, Chughtai KA, Meyer CR, Johnson TD, Galbán S, Rehemtulla A, Kazerooni EA, Martinez FJ, Ross BD. Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nature medicine. 2012; 18:1711–1715.

[30]. Gevenois PA, De Maertelaer V, De Vuyst P, Zanen J, Yernault J-C. Comparison of computed density and macroscopic morphometry in pulmonary emphysema. American journal of respiratory and critical care medicine. 1995; 152(2):653–657. [PubMed: 7633722]

[31]. Coxson HO, Dirksen A, Edwards LD, Yates JC, Agusti A, Bakke P, Calverley P, Celli B, Crim C, Duvoix A, Fauerbach PN, Lomas DA, MacNee W, Mayer RJ, Miller BE, Muller NL, Rennard SI, Silverman EK, Tal-Singer R, Wouters EFM, Vestbo J. The presence and progression of emphysema in COPD as determined by CT scanning and biomarker expression: A prospective analysis from the ECLIPSE study. The Lancet Respiratory Medicine. 2013; 1(2):129–136. [PubMed: 24429093]

[32]. Angelini ED, Delon J, Capelle L, Mandonnet E. Differential mri analysis for quantification of low grade glioma growth. Medical image analysis. 2012; 16(1):114–126. [PubMed: 21911309]

[33]. Mishima M, Hirai T, Itoh H, Nakano Y, Sakai H, Muro S, Nishimura K, Oku Y, Chin K, Ohi M, Nakamura T, Bates JHT, Alencar AM, Suki B. Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. Proceedings of the National Academy of Sciences. 1999; 96(16):8829–8834.

[34]. Blechschmidt RA, Werthschutzky R, Lorcher U. Automated CT image evaluation of the lung: a morphology-based concept. IEEE Transactions on Medical Imaging. 2001; 20(5):434–442. [PubMed: 11403202]

[35]. Achenbach T, Weinheimer O, Buschsieweke C, Heussel CP, Thelen M, Kauczor HU. Fully automatic detection and quantification of emphysema on thin section MD-CT of the chest by a new and dedicated software. RöFo: Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin. 2004; 176(10):1409–15. [PubMed: 15383971]

[36]. Wiemker, R.; Opfer, R.; Bülow, T.; Rogalla, P.; Steinberg, A.; Dharaiya, E.; Subramanyan, K. SPIE Medical Imaging. Vol. 6514. International Society for Optics and Photonics; 2007. Toward computer-aided emphysema quantification on ultra-low-dose CT: reproducibility of ventrodorsal gravity effect measurement and correction.

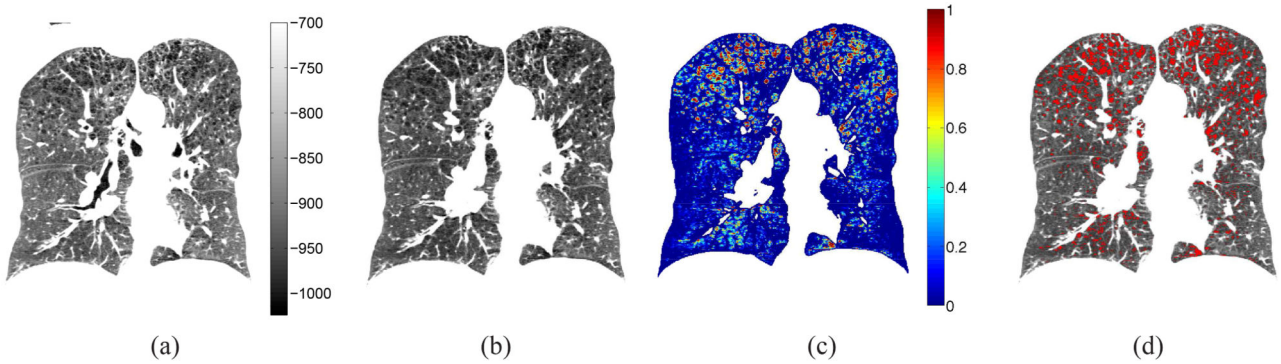


Fig. 1.

Example of the HMMF emphysema segmentation process with corresponding coronal views of: (a) original CT data in the range [−1024, −700] HU, (b) segmented lung region, (c) continuous-valued MAP estimate of the Markov measure field q_1, which represents the emphysema class, and (d) final binary segmentation result f, with red corresponding to the emphysema class, found by maximizing $P(f|q = q^*, \theta = \theta^*, I)$, where q^* and θ^* represent respective MAP estimates of q and θ, given image I.

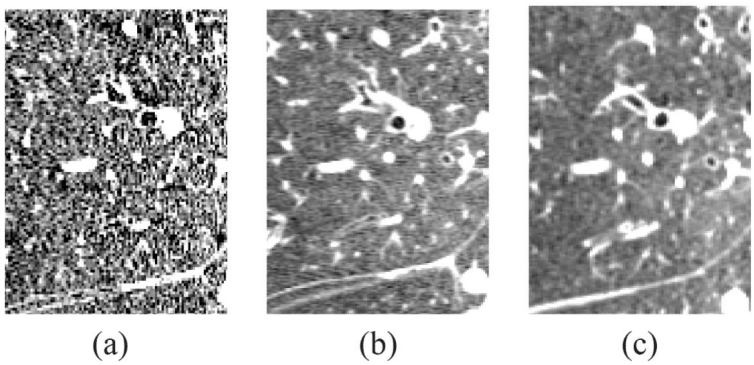


Fig. 2.

Coronal views of a small lung region for a single subject on three CT scans acquired with different imaging protocols: (a) EMCAP B60f (sharp) from 2007, (b) EMCAP B31f (smooth) from 2008, and (c) MESA COPD from 2009.

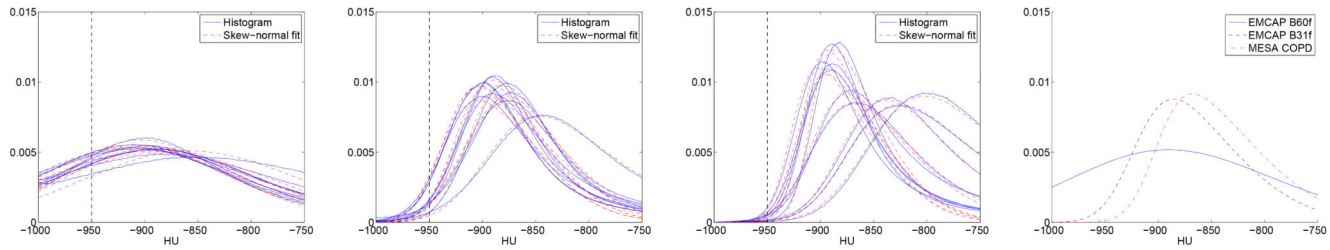


Fig. 3.

(a)-(c) Fitting of skew-normal distributions to normalized intensity histograms of 10 training scans for each imaging protocol. The vertical line indicates the −950 HU threshold used for standard emphysema quantification. (d) Skew-normal distributions with estimated parameter values σ_2^*, α_2^*, and θ_2^0 for the three imaging protocols (see Table II).

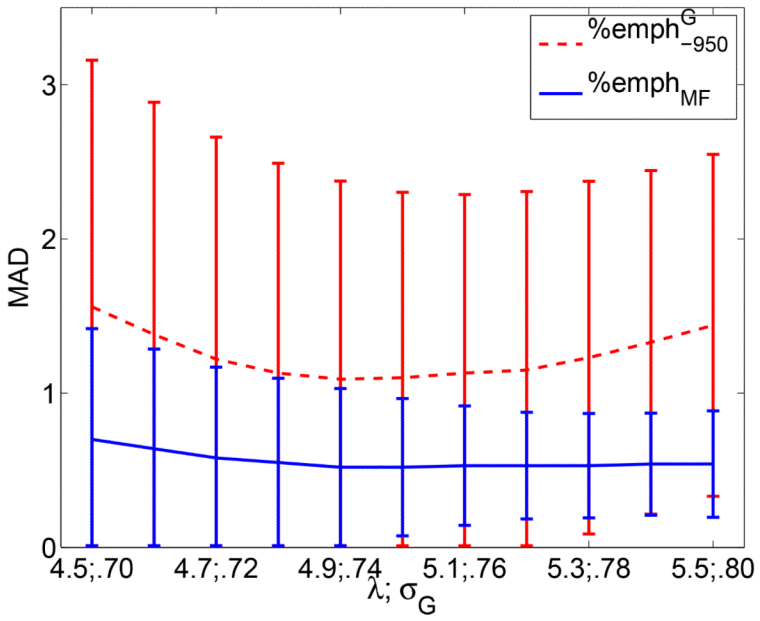


Fig. 4.

Mean absolute differences (MAD) and standard deviations of $\%emph^{G}_{-950}$ and %*emph*$_{\mathrm{MF}}$ in the parameter training set, as a function of σ_G and λ, respectively.

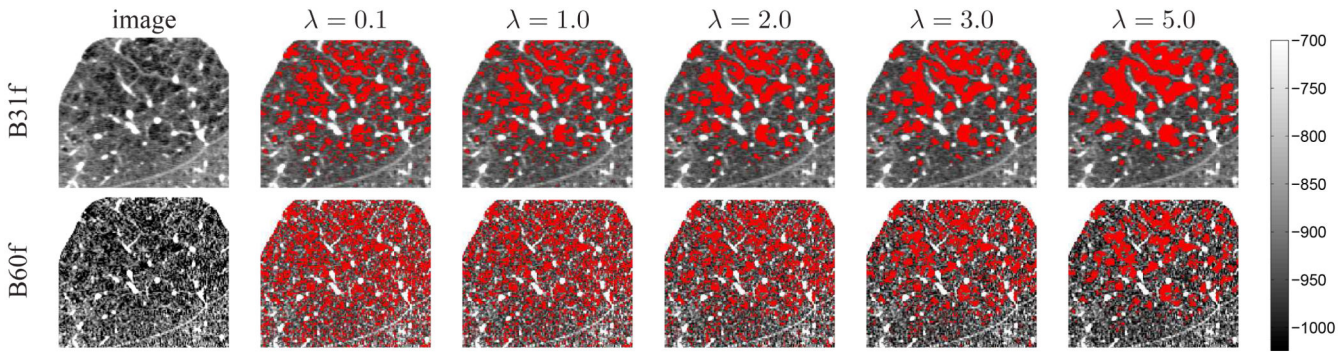
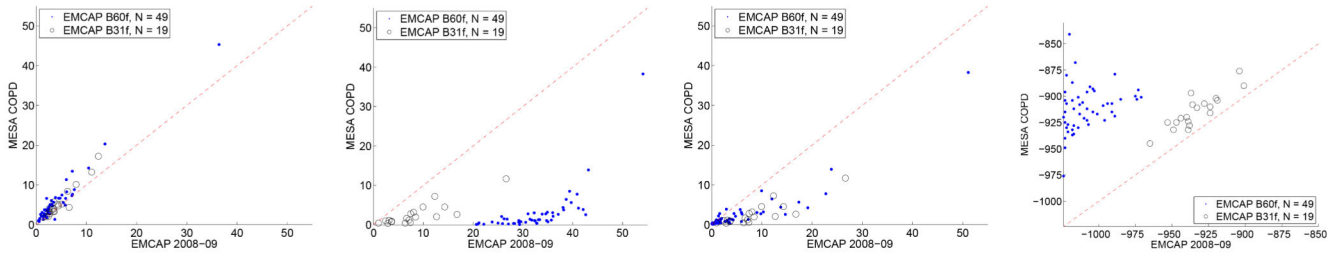


Fig. 5.

Illustration of the HMMF segmentation results for different values of the Markov field weight λ. The images represent a cropped coronal view of a scan reconstructed with the two different kernels, with B31f reconstruction on the top row and B60f reconstruction on the bottom row. The red color represents regions classified as emphysema.



(a) $\%emph_{\mathrm{MF}}$ (b) $\%emph_{-950}$ (c) $\%emph^{G}_{-950}$ (d) PD_{15} (HU)

Fig. 6.

Scatterplots of emphysema scores between EMCAP scans acquired in 2008-09 and MESA COPD scans. The total number of scans was 68, of which 19 were reconstructed with the B31f kernel in EMCAP. Diagonal line represents one-to-one correspondence.

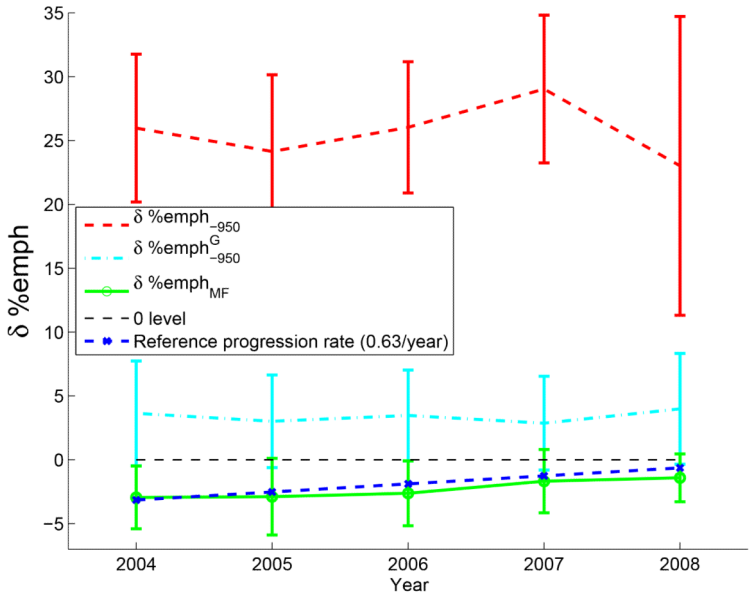
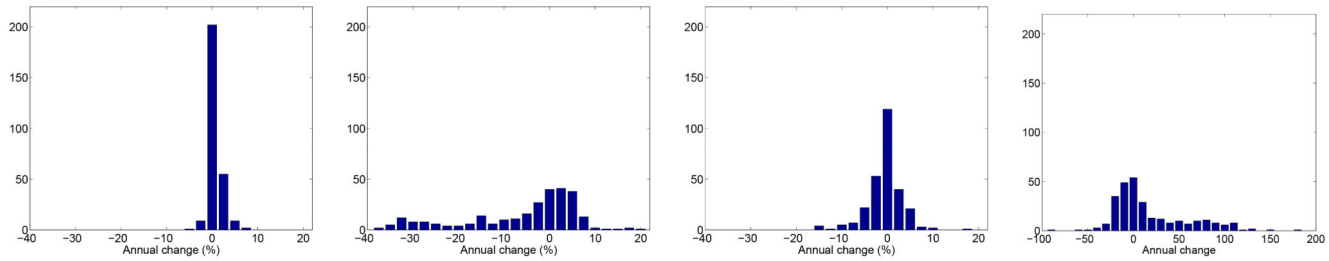


Fig. 7.

Means and standard deviations of differential emphysema measures $\delta\%emph$ between EMCAP (2004 – 2008) and MESA COPD (2009 – 2011). Number of scans for each year is reported in Table I. A reference annual progression rate of 0.63 for $\%emph_{-950}$ [31] is plotted in blue, by assigning a value of $-0.63 \cdot (2009 - Y)$ for each year Y.



(a) $\Delta_a \%emph_{\mathrm{MF}}$ (b) $\Delta_a \%emph_{-950}$ (c) $\Delta_a \%emph^{G}_{-950}$ (d) $\Delta_a PD_{15}$

Fig. 8.

Histograms of 278 evaluations of annual changes for the four emphysema measures. Please note the different x-axis for the PD_{15} measure in (d).

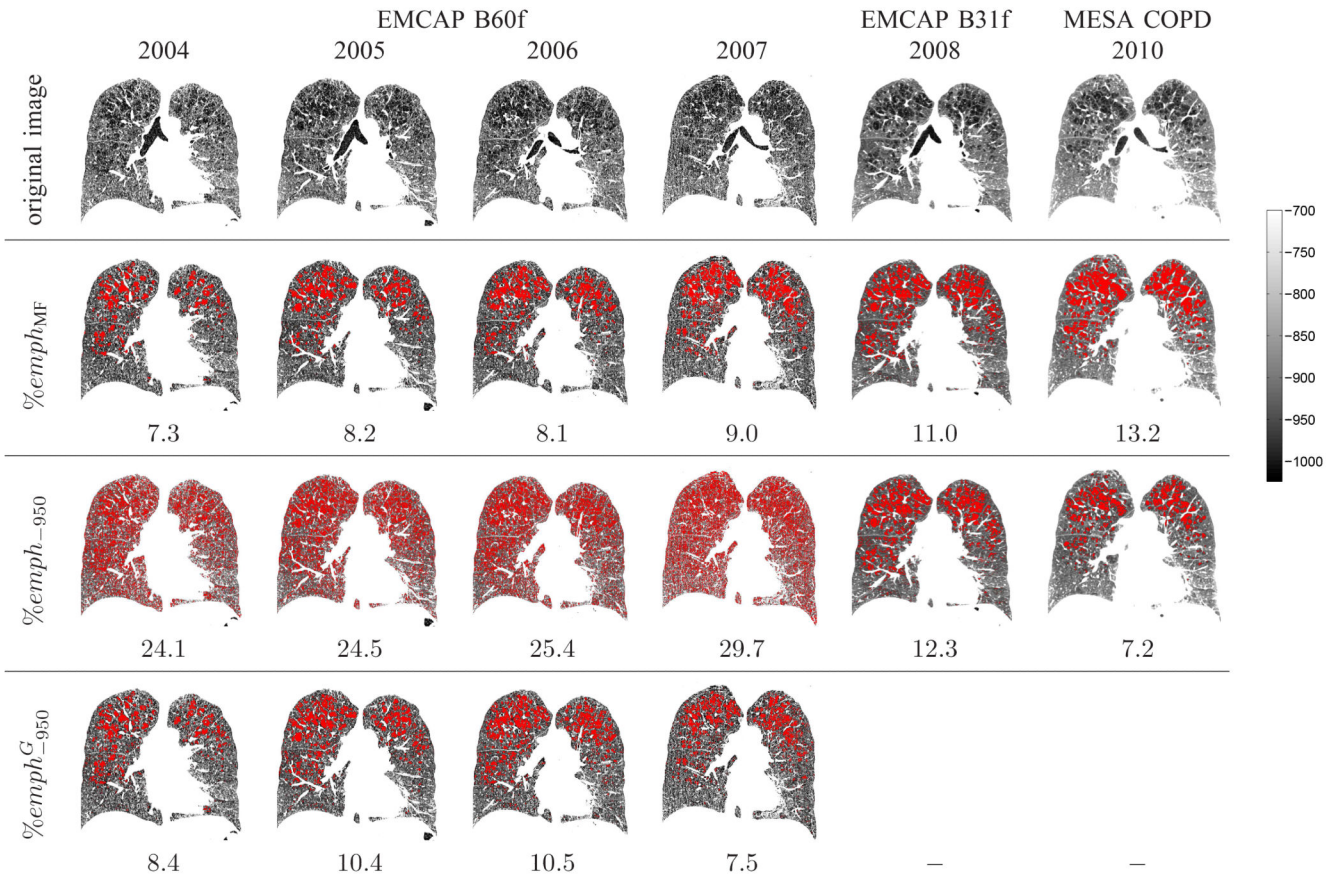


Fig. 9.

Example of emphysema masks and %*emph* values for six scans for a single subject, between the years 2004 and 2010. The top row shows original image slices, and the three bottom rows show emphysema masks generated by the HMMF method (%$emph_{\mathrm{MF}}$), thresholding at −950 HU (%$emph_{-950}$), and thresholding at −950 HU with prior Gaussian smoothing ($\%emph^{G}_{-950}$). Scans from 2004 to 2008 were from the EMCAP study, and the latest scan was from MESA COPD. Scans between 2004 and 2007 were reconstructed with the B60f kernel and the last EMCAP scan with the B31f kernel.

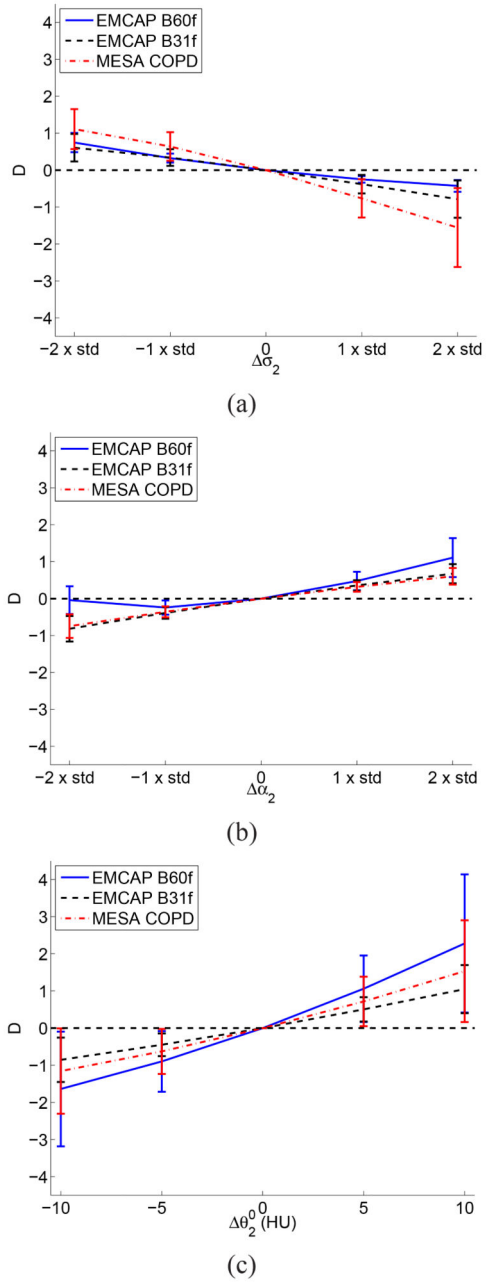


Fig. 10.

Mean differences (D) in the %$emph_{MF}$ scores, resulting from modifying the values of the parameters (a) σ_2 by $\Delta\sigma_2$, (b) α_2 by $\Delta\alpha_2$, and (c) θ_2^0 by $\Delta\theta_2^0$. The notation 'std' refers to the standard deviations of the estimated parameter values reported in Table II. The errorbars represent standard deviations of D, evaluated over the 22 EMCAP B60f and B31f training scans, and a randomly selected set of 20 scans from MESA COPD.

TABLE I

Number of scans for each year for the 87 subjects in the evaluation data set. Each subject had no more than one scan each year. Of the last EMCAP scans (2008-2009), 6 were acquired in early 2009, with the remainder in 2008. Of the 87 MESA COPD scans, 55 were acquired in 2009, 27 in 2010 and 5 in 2011. The total number of scans was 365.

	Year	Siemens B60f	Siemens B31f	GE
Total	**2004 – 2011**	**259**	**19**	**87**
MESA COPD	2009 – 2011			87
EMCAP	2008 – 2009	49	19	
	2007	73		
	2006	72		
	2005	42		
	2004	23		

TABLE II

Estimated imaging protocol-dependent parameter values used in the evaluation study, and descriptions of the training data used for their estimation (details of the data are presented in the text): mean values and standard deviations of σ_2^* and α_2^* of skew-normal distributions for class 2 (parenchyma), location $\mu_{\theta_1}^* \left(=\theta_1^0\right)$ for class 1 (emphysema) prior distribution and the associated tracheal air intensity I_{tr}^*, initial value θ_2^0, and the Markov field weight λ.

	σ_2^* (HU)	α_2^*	$\mu_{\theta_1}^* \left(=\theta_1^0\right)$ (HU)	θ_2^0 (HU)	λ
CT scans used for parameter value estimation	**Low *%emph*$_{-950}$**	**Low *%emph*$_{-950}$**	**Any**	**Any**	**Multiple reconstructions**
EMCAP B60f	124.4(±6.3)	1.25(±0.26)	−1000 (I_{tr}^* = −954)	−982	5.0
EMCAP B31f	76.9(±10.5)	3.01(±0.42)	−1028 (I_{tr}^* = −983)	−949	1.0
MESA COPD	79.5(±17.7)	3.74(±0.51)	−1000 (I_{tr}^* = −931)	−931	1.0

TABLE III

Means, standard deviations, and minimum and maximum values of the emphysema scores over the entire data set. 1) EMCAP B60f, all (2004 – 2009); 2) EMCAP B60f, most recent (2006 – 2009); 3) EMCAP B31f (2008 – 2009); 4) MESA COPD, subjects in EMCAP B31f (2009 – 2011); 5) MESA COPD, all (2009 – 2011)

Data set	N subjects	N scans	$\%emph_{MF}$ mean ± std., [min,max]	$\%emph_{-950}$ mean ± std., [min,max]	$\%emph^{G}_{-950}$ mean ± std., [min,max]	PD_{15} mean ± std., [min,max]
1	87	259	3.4 ± 4.4, [0.2, 36.4]	30.1 ± 7.2, [10.3, 54.1]	6.0 ± 7.6, [0.0, 51.1]	−998 ± 19, [−1024,−922]
2	87	87	3.8 ± 4.5, [0.4,36.4]	32.4 ± 6.3, [18.5, 54.1]	6.2 ± 7.7, [0.0, 51.1]	−1007 ± 15, [−1024, −967]
3	19	19	4.7 ± 3.0, [1.4,12.4]	8.5 ± 6.1, [1.0, 26.6]	-	−934 ± 16, [−965, −901]
4	19	19	5.5 ± 4.1, [1.8,17.2]	2.6 ± 2.8, [0.3,11.7]	-	−914 ± 16, [−945, −876]
5	87	87	5.5 ± 5.9, [0.7,45.3]	2.8 ± 4.7, [0.1, 38.3]	-	−912 ± 21, [−976, −841]

TABLE IV

Pairwise correlations and 95% confidence intervals (in square brackets) of emphysema measures between longitudinal scans. Reported values are statistically significant ($p < 0.0001$), except when marked with '*n*'. Highest correlation of each comparison is shown in bold. t_1 and t_2 indicate the set of data being used, and *N* refers to the number of scans in each comparison. For space considerations, comparisons with less than 16 cases were omitted, and the following shorthand expressions were used: MESA COPD (MC), EMCAP B31f (B31f), EMCAP B60f 2008-09 ('08) and other B60f years accordingly, $\%emph_{MF}$ ($\%_{MF}$) and other *%emph* measures similarly, PD_{15} (*PD*).

t_1	MC	MC	MC	MC	MC	MC	B31f	B31f
t_2	B31f	'08	'07	'06	'05	'04	'06	'05
N	**19**	**49**	**73**	**72**	**42**	**23**	**18**	**16**
$\%_{MF}$	0.96[0.89, 0.98]	0.98[0.96, 0.99]	0.95[0.92, 0.97]	0.97[0.95, 0.98]	0.99[0.97, 0.99]	0.96[0.90, 0.98]	0.88[0.69, 0.95]	0.90[0.74, 0.97]
$\%_{-950}$	0.86[0.66,0.94]	0.70[0.52, 0.82]	0.52[0.33, 0.67]	0.65[0.50, 0.77]	0.67[0.47,0.81]	0.61^n [0.27, 0.82]	0.69^n [0.33, 0.87]	0.64^n [0.22, 0.86]
$\%^G_{-950}$	—	0.91[0.85, 0.95]	0.86[0.78, 0.91]	0.90[0.85, 0.94]	0.93[0.87, 0.96]	0.88[0.73, 0.95]	0.85[0.63, 0.94]	0.78^n [0.46, 0.92]
PD	0.87[0.70,0.95]	0.26^n [−0.02, 0.51]	0.38^n [0.17, 0.56]	0.56[0.38, 0.70]	0.60[0.37, 0.77]	0.62n [0.28, 0.82]	0.58n[0.16, 0.83]	0.68n [0.27, 0.88]
t_1	'08	'08	'07	'07	'07	'06	'06	'05
t_2	'07	'06	'06	'05	'04	'05	'04	'04
N	42	39	61	37	18	36	17	17
$\%_{MF}$	0.96[0.93, 0.98]	0.98[0.96, 0.99]	0.96[0.93, 0.97]	0.98[0.97, 0.99]	0.85[0.63,0.94]	0.98[0.96, 0.99]	0.95[0.88, 0.98]	0.99[0.96, 1.00]
$\%_{-950}$	0.73[0.55,0.85]	0.84[0.71, 0.91]	0.76[0.63, 0.85]	0.78[0.61, 0.88]	0.84[0.60,0.94]	0.86[0.74, 0.93]	0.89[0.72, 0.96]	0.87[0.68, 0.95]
$\%^G_{-950}$	0.88[0.79,0.93]	0.91[0.84, 0.95]	0.89[0.82, 0.93]	0.91[0.83, 0.95]	0.72^n [0.39, 0.89]	0.95[0.91, 0.98]	0.94[0.84, 0.98]	0.95[0.87, 0.98]
PD	0.50^n [0.24, 0.70]	0.74 [0.55, 0.85]	0.63[0.45, 0.76]	0.68[0.46, 0.82]	0.86[0.65, 0.95]	0.83[0.68, 0.91]	0.88[0.70, 0.96]	0.90[0.75, 0.96]

TABLE V

Pairwise correlations between $\%emph_{MF}$ and $\%emph_{-950}$, with all p-values < 0.0001, except for the year 2004 ($p < 0.01$).

	Year	Correlation	Number of scans
MESA COPD	2009 – 11	0.98	$n = 87$
EMCAP B31f	2008 – 09	0.79	$n = 19$
EMCAP B60f	2008 – 09	0.74	$n = 49$
	2007	0.60	$n = 73$
	2006	0.70	$n = 72$
	2005	0.64	$n = 42$
	2004	(0.57)	$n = 23$