Journal of Digital Imaging

# Lung Field Segmenting in Dual-Energy Subtraction Chest X-ray Images

Robert E. Alvarez, Ph.D.

**The purpose of this study was to develop and test a method to delineate lung field boundaries in dual-energy chest x-ray images. The segmenting method uses soft-tissue images and spatial frequency–dependent, background-subtracted images. Large-scale chest anatomy features are located and used to select the lung apices, the lateral lung boundaries, and the lung–mediastinum and lung–diaphragm boundaries. Extraneous parts of the contours are removed and they are joined to form complete lung boundaries. The reliability measure uses a statistical shape model to estimate the probability of occurrence of a contour. The method was experimentally tested with 30 human subject images. It has higher accuracy and specificity and a sensitivity parameter equal to the best previously reported method. The reliability measure is able to detect contours with unusual lung outlines or errors in the processing. The method exploits the characteristics of dual-energy subtraction images to improve lung field segmenting performance.**

**KEY WORDS: Lung, segmenting, dual-energy subtraction, chest, x-ray**

LUNG FIELD SEGMENTING, the delineation of the boundaries of the lungs in a chest x-ray image, is an essential part of a number of automated chest x-ray image analysis systems. The lung outlines can be used directly as indicators of cardiomegaly[1] or pneumothorax,[2] and they are also used in many computer-assisted detection or diagnosis (CAD) systems[3] to determine the area of processing. The soft-tissue images[4] provided by dual-energy technology[5] eliminate ribs, which are a major source of errors in these CAD systems. Previous work in segmenting dual-energy soft-tissue images[6] used algorithms developed for conventional images. In this article, I describe a segmenting method that exploits the characteristics of dual-energy soft-tissue images to improve performance.

The performance of the segmenting method is tested by applying it to a set of human subject dual-energy images. The performance is measured by computing the accuracy, sensitivity, and specificity using the approach introduced by Duryea and Boone,[7] which was also used by Vittitoe et al.[8] and van Ginneken and Ter Haar Romeny.[9] Van Ginneken and Ter Haar Romeny compute the performance measures for other lung field segmenting methods reported in the literature and provide a table of their values. The results presented here show that the dual-energy soft-tissue image segmenting method has higher accuracy and specificity and a sensitivity parameter equal to the best previously reported method.

An essential element of any automated medical image analysis system is an estimate of the reliability of the results. Biological variability can result in degraded performance, and an indication of reliability allows for corrective action or, at the least, warning the user of potential problems. I describe a reliability measure for lung field segmenting algorithms based on statistical shape models.[10] The measure uses a training set of manually digitized lung boundaries and principal components analysis to compute an estimate of the probability of occurrence of a contour produced by an automated algorithm. Evaluation of the reliability measure with

*From the Aprend Technology, 2369 Laura Lane, Mountain View, 94043, California.*

*Correspondence to: Robert E. Alvarez, Ph.D., tel.: 650-968-9794; fax: 208-247-7469; e-mail: JDISE@APRENDTECH.com*



*Online publication 2 March 2004*
*doi: 10.1007/s10278-003-1701-8*

the contours produced by the dual-energy soft-tissue image segmenting method shows that it is able to detect both errors in the processing and unusual lung anatomy.

Segmenting methods for conventional chest x-ray images were reviewed by van Ginneken (Ch. 2).[11] The methods use two general approaches: rules-based and pixel classification. Rules-based methods apply a sequence of image processing operations such as image linear combination, spatial filtering, thresholding, and morphological operations to delineate the lung field boundaries. In pixel classification, each pixel in the image is individually classified as lung/non-lung based on features such as image data magnitude, location, and local texture measures. Either approach can utilize global knowledge, such as human thorax anatomy, (1) implicitly through constraints and tests in the algorithm logic, or (2) explicitly through a series of rules expressed as logical constraints,[12] or with statistical shape models trained on hand-digitized contours.[13]

As noted by Vittitoe et al.,[8] the lung boundaries derived by the previous methods and the method described in this article are actually the unobscured lung regions—that is, those regions of the lungs not obscured by the mediastinum, heart, or diaphragm. The unobscured region has been estimated to represent only approximately 57% of the lungs,[14] so methods to segment and process the remaining parts of the lungs are important. Unfortunately, segmenting the complete area of both lungs is substantially more difficult than segmenting the unobscured regions and it is the subject of ongoing research.

## MATERIALS AND METHODS

### Segmenting Method Description

The method has four principal steps:

1. Locate large-scale chest anatomy features: the thorax centerline, and the centerlines of the lungs.
2. Determine lung boundary contours as edges in spatial derivatives of the soft-tissue and background-subtracted images.
3. Eliminate extraneous parts of the contours that are not part of the lung boundary with rules-based pruning methods.
4. Resample the contours so that points correspond to those of the statistical shape model.

### Image Contour Computation and Selection

The lung boundary contours are computed from local maxima (or minima) in spatial derivatives of the image data with an approach similar to that used in the "rules-based reasoning method" of van Ginneken and Ter Haar Romeny.[9] The spatial derivative images are computed by convolving the original image with spatial derivatives of a two-dimensional Gaussian function. By taking the two-dimensional Fourier transform, it is easy to see that this method is equivalent to smoothing the derivative of the image with a Gaussian function. This approach is less sensitive to noise and more suited to relatively large features than alternatives such as Robert's cross or Sobel operators based on local differences.

We convolve the images using a standard algorithm that extends the image values outside the original boundaries by mirror reflection across the edge. The Gaussian functions used as convolution kernels are truncated at a specified multiplier of their standard deviations. We compute gradient magnitude images by combining the $x$ and $y$ derivative images pixel by pixel $I_{GradientMagnitude}(x,y) = \sqrt{I^2_{xderiv}(x,y) + I^2_{yderiv}(x,y)}$.

A set of contours is computed from the local maxima images by finding sets of 8-connected points using standard computer vision software. Contours that fit selected geometric criteria are selected by a voting process.[9,15] The method starts by associating a count variable with every contour in an image and initializing it to zero. It raises the count incrementally every time the contour satisfies a criterion. An example is to find the contour closest to a line $x = x_0$ in a rectangular region. For each $y$ value in the region, the method searches along the $x$-direction to find the closest contour passing through or contained within the region and raises its vote count. After searching for all $y$-values in the range, the contour with the highest number of votes is selected. As an option, other contours whose $y$ values do not overlap with the high vote contour can be included in the final result.

### Background-Subtracted Soft-Tissue Image

In addition to the soft-tissue image, we use a special background-subtracted image as input to the contour computation. The background-subtracted image is created by subtracting a frequency-dependent smoothed version of the soft-tissue image. Unlike unsharpness masking, this method preserves edges in the image, because the amount of smoothing is smaller when the magnitude of the gradient of the original image is larger as at edges. The image is computed by first creating three images using a low, medium, and high degree of smoothing. A gradient magnitude image is also computed and it is scaled so its values are from 0 to 1. The final image is computed by combining, at each pixel, the values from the low to high smoothed images depending on the magnitude of the gradient at that pixel. The combination functions are shown in Figure 1. An example of a background-subtracted image is shown in Figure 2.

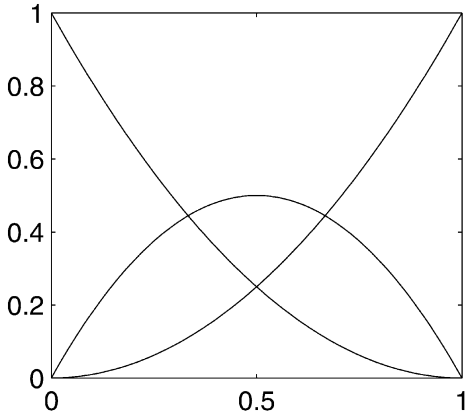


**Fig 1. Combination functions for computing the frequency-dependent smoothed image. The independent variable (*x*-axis) is the normalized magnitude of the gradient. The three functions shown multiply the image data from the low (maximum value at 0), medium (maximum value at 0.5), and highly smoothed (maximum value at 1) images.**

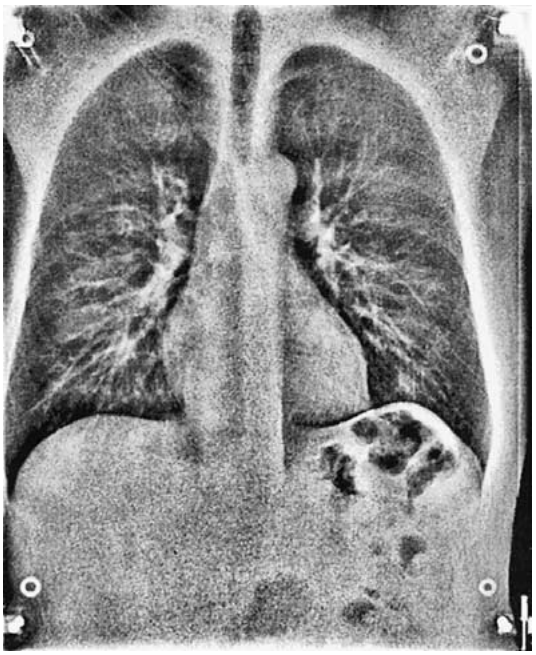

**Fig 2. Example of a background-subtracted soft-tissue image used for contour computation. Background variation is eliminated while retaining edges by subtraction of a frequency-dependent smoothed image.**

## Locate Large-Scale Anatomy Features

The large-scale anatomy features are computed from a smoothed soft-tissue image using a Gaussian standard deviation of 0.05 (all dimensions are fractions of the number of pixels in the horizontal *x*-dimension of the image). The thorax centerline is selected from the contours resulting from local minima in the *x*-direction. The selection is done by voting for the contours closest to the line $x = 0.5$ in a rectangular region with corners at (0.3,0.2) and (0.7,0.6). After the thorax centerline is determined, its average *x*-value, $x_{cl}$, is computed and used in the further processing. The lung centerlines are selected from contours resulting from local maxima in the *x*-direction. The right lung centerline is computed by voting for contours closest to the line $x = 0.25$ in a box with corners (0.03, 0.2) and ($x_{cl}$, 0.6). The left lung centerline results from contours closest to $x = 0.75$ in a box with corners ($x_{cl}$, 0.2) and (0.97,0.6). Once the centerline contours were determined, straight lines were fit to each contour for use in later processing. In addition, lines perpendicular to the lung centerlines through the lung contour center of mass were computed. Figure 3 shows an image with centerlines, best-fit lines, and the lung boundary contours computed as described in the paragraphs that follow.

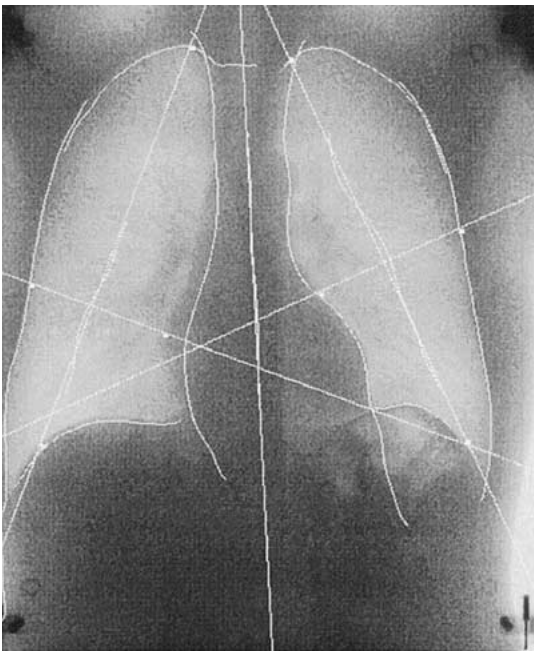

**Fig 3. Large-scale thorax anatomy features: thorax and lung centerlines. The image also shows the lung contours before the pruning step. The contours are the bright white lines. See text for description of computation method.**

## Locate Lung Apices and Diaphragm Boundaries

First, individual points on the apices and diaphragms were computed by finding the maximum values of the gradient of the background-subtracted image on the straight lines through the lung centerline contours. Next, the full contours were computed using the voting method in local maxima of the soft-tissue gradient-magnitude image. The

right lung apex and diaphragm were the closest contours in the $y$-direction to the $y$-coordinate of the intersection points. The voting process was limited to a 0.025 range of $x$-coordinates centered on the $x$-coordinate of the intersection point. The left lung apex $y$-value was unreliable, so the right lung apex value was used in the contour selection process, relying on the symmetry of the lungs. The $x$-range of the search was centered on the intersection of the left lung centerline with the horizontal line $y = y_{rightapex}$.

In a number of cases, the clavicle was not completely subtracted in the soft-tissue image, causing an error in the apex determination. To correct those cases, a rule was added that if the $y$-coordinate of the apex point was greater than 0.2 (the coordinate system is inverted so increasing $y$ coordinates are toward the bottom of the image), the search was limited to the region beyond that point, In addition, in some cases, the diaphragm maximum was found on a small contour resulting from lung structure. These cases were rejected by testing the size of the contour. If it was less than 0.1, the maxima were located on a vertical line passing through the center of mass of the lung centerline.

## Lung-Mediastinum Boundary Contours

The lung–mediastinum contours were computed from the local maxima (right) or minima (left) of the $x$-derivative of the background-subtracted image. A point on the contour was determined as the maximum (right) or minimum (left) value along the line perpendicular to the lung centerline in the segment from the lung centerline contour center of mass to the intersection with the thorax centerline. Next, the contours were determined by voting for contours closest in the $x$-direction in a 0.025 $x$-range centered on the intersection points and from 0.2 to 0,4 in the $y$-direction.

## Lateral Lung Boundaries

The lateral (i.e., farthest from the thorax center also called "rib cage" boundary) lung contours were computed from the local maxima (left) or minima (right) of the $x$-derivative of the soft-tissue image. A point on the contour was determined as the maximum (left) or minimum (right) along the perpendicular to the lung centerline on a segment from the outside edge of the image to three quarters of the distance to the center of mass of the lung centerline contour. Then the lateral lung boundary contours were selected as the closest contours to the intersection point $x$-value along the $x$-direction in a square region centered on the intersection point with sides equal to 0.04.

## Prune Contours

The complete contour of each lung was formed from four contours: the apex, mediastinum, diaphragm, and the lateral lung boundaries, as shown in Figure 3. In some cases, the contours extended past the lung, and in other cases they covered only part of the true lung boundary. The final lung field contours were computed in these cases with a "pruning" procedure. The procedure either removed the parts of the contours extending beyond the intersections of the individual pieces or extended the ends of the pieces that did not intersect with tangential straight lines fitted to the end parts of the contours.

An algorithm was used to form complete lung contours from the individual parts. On input, each component contour's points were sorted. The apex and diaphragm boundaries were sorted in order of increasing $x$-coordinate, and the mediastinum and lateral boundaries were sorted in order of increasing $y$-coordinate. Two binary parameters, called prune_before1 and prune_before2 in the listing below, were used to indicate whether to remove or extend the part of the contour before or after the intersection.

The ProcessCorner(contour1,prune_before1,contour2, prune_before2) algorithm description in pseudo-code is as follows:

```
1. intersect_point = FindIntersect(contour1, contour2)
2. if (contours intersect) then begin
3. ClipContour(contour1, intersect_point, prune_before1)
4. ClipContour(contour2, intersect_point, prune_before2)
5. end
6. else begin
7. line1 = FitTangentLine(contour1, prune_before1)
8. line2 = FitTangentLine(contour2, prune_before2)
9. intersect_point = FindIntersect(line1, line2)
10. point1 = FindNearestPoint(contour1, intersect_point)
11. point2 = FindNearestPoint(contour2, intersect_point)
12. ClipContour(contour1, point1, prune_before1)
13. ClipContour(contour2, point2, prune_before2)
14. end
```

The FindIntersect(contour1, contour2) procedure finds the actual intersection point of the two contours or fits tangential straight lines to the half of each contour closest to the intersection and finds their intersection. The ClipContour(contour, point, prune_before) procedure removes points from the contour before or after the specified point depending on whether the parameter is **true** or **false**. The FitTangentLine(contour, prune_before) procedure fits a straight line to half of the points at the beginning or end of the contour, depending on the value of the prune_before parameter.

The ProcessCorner procedure was used repeatedly to join the right lung contour pieces with the following parameters:

1. ProcessCorner(right_apex, false, right_lateral, false)
2. ProcessCorner(right_diaphragm, true, right_mediastinum, true)
3. ProcessCorner(right_apex, true, right_mediastinum, false)
4. ProcessCorner(right_diaphragm, false, right_lateral, true)

Similarly, the left lung contours were processed as follows:

1. ProcessCorner(left_apex, false, left_mediastinum, false)
2. ProcessCorner(left_mediastinum, true, left_diaphragm, false)
3. ProcessCorner(left_diaphragm, true, left_lateral, true)
4. ProcessCorner(left_apex, true, left_lateral, false)

The values of the prune_before parameters were determined by studying a diagram of the lung boundary and using the

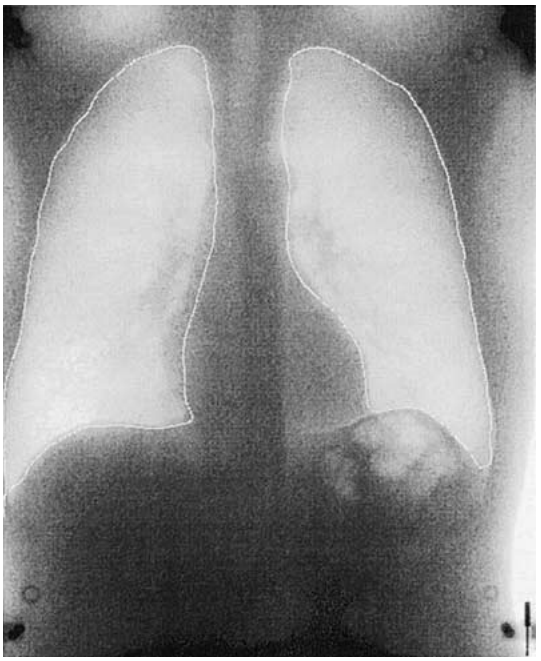

**Fig 4. Final lung boundaries after the pruning process. This is the output of the third step in the segmenting method.**

order of sorting of the contour points described above. An example of the complete boundaries for both lungs is shown in Figure 4.

## Resample the Contours

The statistical shape model uses training data that are manually digitized. The training data consist of 64 points per lung, 128 points in all. The computer-generated contours are sampled at high density, essentially one point per pixel. The contours were resampled to correspond to the manually entered contours by locating landmark points, shown in Figure 5, using geometric constraints. Equally spaced points are inserted between the landmarks using spline interpolation. A procedure, NearestPoint(contour, point), finds the contour point closest to the specified point by exhaustive search. The Point(y) method finds the contour point with the specified $y$-value. These procedures were applied to find the landmark points as shown by the following pseudo-code:

Right lung contour:

1. point = (right_contour.xmin, right_contour.ymax)
2. right_contour_point_0 = NearestPoint(right_contour, point)
3. right_contour_point_28 = Point(right_contour.ymin)
4. point = (right_contour.xmax, right_contour.ymax)
5. right_contour_point_51 = NearestPoint(right_contour, point)

Left lung contour

1. point = (left_contour.xmax, left_contour.ymax)
2. left_contour_point_0 = NearestPoint(left_contour, point)

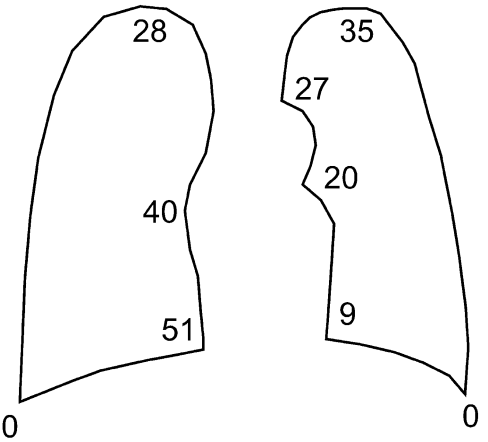


**Fig 5. Landmark points used for manual entry of lung contours.**

3. point = (left_contour.xmin, left_contour.ymax)
4. left_contour_point_9 = NearestPoint(left_contour, point)
5. left_contour_point_35 = Point(left_contour.ymin)

After the landmark points were found, other points were inserted, equally spaced, along the original contour by using spline interpolation of the $x$- and $y$-coordinates.

## Statistical Shape Model

Before the results of any automatic processing are used, they should be tested to estimate their reliability. The variability of biological structures makes this task difficult. Although there is obvious regularity in the structure, the variability implies the need for a statistical model, and the results of a reliability test will be expressed as probabilities. The statistical shape model introduced by Cootes et al.[10] was used to develop a segmenting method reliability measure that uses prior knowledge of regularities in human anatomy.

The statistical shape model starts with a set of landmark points, i.e., points on the lung contour that can be consistently located in different examples of the object. Suitable points are at corners, junctions, and regions of high curvature. To describe the shape more accurately, additional points are inserted, equally spaced, between the major landmark points. The shape is represented by concatenating the coordinates of the individual landmark points in the order of a traversal of the contour. A clockwise traversal was used for the segmenting method. If there are $p$ landmark points, the shape is represented by a single $2p$ element vector

$$x = (x_1, y_1, x_2, y_2, \ldots, x_p, y_p). \quad (1)$$

Note that, unlike the model of Cootes et al. a row vector is used to represent the coordinates. This is consistent with principal component analysis literature and corresponds to the representation used in the computations.

The coordinates in Eq 1 are subjected to an "alignment" procedure that removes location, scale, and rotational

effects. For the lung contours, each coordinate was divided by the number of pixels in the $x$-direction. The contours for both lungs were combined into a single shape vector and then translated so that the center of mass was at the origin. They were then scaled so that the range of $y$-coordinates was equal to one. The coordinates were not rotated, because patients were aligned during the x-ray examination by the radiological technologist.

The statistical shape model is derived from a set of $t$ training images, usually manually digitized by an experienced observer. Then, principal component analysis[16] is used to create a parameterized model of the training data with smaller dimensionality. The principal components are computed as follows:

1. Compute the mean of the data;

$$\bar{x} = \frac{1}{t}\sum_{i=1}^{t} x_i \tag{2}$$

2. Compute the covariance as follows (the superscript $T$ denotes a transpose);

$$S = \frac{1}{t-1}\sum_{i=1}^{t} (x_i - \bar{x})^T(\bar{x}_i - \bar{x}) \tag{3}$$

3. Compute the eigenvectors, $\phi i$, and the corresponding eigenvalues, $\alpha_i$.;

$$S_\phi = \lambda\phi \tag{4}$$

The eigenvectors specify a transformation of the original data into a new set of uncorrelated variables, called the *principal components*. The eigenvalues, also called *characteristic* or *latent roots*, are the variances of the new variables along their respective axes. Since the sum of the eigenvalues is equal to the sum of the variances of the original variables, the eigenvalues represent the part of the total variance represented by each of the principal components.

In many cases, a relatively small number of principal components can be used to approximate the original data while still representing a large fraction of its variation. This is done by sorting the eigenvalues and corresponding eigenvectors in order of decreasing magnitude: $\alpha_1 > \alpha_2 > \ldots$ A number of modes used is selected so that the model represents a large fraction of the total variance of the training data. The corresponding eigenvectors allow us to approximate any member of the original data set using an $m$ dimensional vector, **b**,

$$\hat{x} = \overline{x} + b\Phi^T \tag{5}$$

where $\phi = (\phi_1\phi_2|\ldots|\phi_m^T$ and $b = (x - \overline{x})\Phi$. The results shown later indicate that a small number of modes, compared to the number of training images or the dimensionality of the shape vectors, can approximate the lung contour shapes.

## RELIABILITY MEASURES: STANDARDIZED RESIDUALS AND SCALE PARAMETER

To test the reliability of a shape, we can compute the residuals, i.e., the differences between the shape and its approximation, using an $m$-mode statistical shape model. Jackson and Mudholkar[17] show that, under the assumption that the model parameters, in our case the coordinates, are normally distributed, the function $s(Q)$ of the sum of the squares of the residuals $Q = (x - \hat{x})(x - \hat{x})^T$ is normally distributed with mean 0 and variance 1.

$$S(Q) = \theta_1 \frac{\left[\left(\frac{Q}{\theta_1}\right)^{h_0} - \frac{\theta_2 h_0(h_0-1)}{\theta_1^2} - 1\right]}{\sqrt{2\theta_2 h_0^2}} \tag{6}$$

The constants in the expression depend on the eigenvalues and the number of modes, $m$, of the statistical shape model retained:

$$\theta_1 = \sum_{i=m+1}^{p} \lambda_i, \quad \theta_2 = \sum_{i=m+1}^{p} \lambda_i^2, \quad \theta_3 = \sum_{i=m+1}^{p} \lambda_i^3,$$
$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}.$$

I call $s(Q)$ the *standardized residuals*.

The scale parameter, $M$, computed during the contour alignment procedure, provides another useful measure of a contour independent of shape. Recall that shapes ignore location, scale, and rotational effects. The location and rotation of the patient are controlled by the radiological technologist during the x-ray examination, but the scale depends on the size of the lungs and is subject to biological variability. It is therefore a source of potential errors by the segmenting algorithm.

The reliability measure, $R$, is defined to be the probability of occurrence of a lung contour. If the probability is low, then the results may have lower reliability. The measure is computed as:

$$R(q, m) = \Pr[S_Q > |q|, M > |m|] = \Pr[S_Q > |q|]\Pr[M > |m|],$$

where $S_Q$ is the random variable with the distribution of $S(Q)$ and $M$ is the scale parameter. The second part of Eq 7 is justified because the results below show that $S_Q$ and $M$ are independent. The results also show that they can be approximated as Gaussian random variables, so we can compute the probabilities on the right-hand side of the equation as:

$$\Pr[X > |x|] = 1 - cdf_{normal}(x) + cdf_{normal}(-x) \tag{8}$$

where $cdf_{normal}(x)$ is the cumulative distribution function of a Gaussian random variable with mean $\mu$ and standard deviation $\sigma$

$$cdf_{normal}(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{K} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \tag{9}$$

## Experimental Method

A set of 30 human subject images was used to test the segmenting method and the reliability measure. Ten of the images were acquired using an experimental electro-optical detector and x-ray tube voltage switching with standard computed radiography plates.[18] Twenty images were acquired with a commercial system that uses voltage switching with rapid readout of a solid-state detector (GE Medical

Systems Revolution™ XR/d System). These systems acquire two images with a low and high x-ray tube voltage then process them to compute the soft-tissue image. Each system used different software to compute the soft-tissue images, but the images from both systems can be used as input to the dual-energy segmenting algorithm.

The only inputs to the segmenting software are the soft-tissue image and a parameter indicating the contrast of the data. As shown in the results, the two systems have differing contrast—i.e., thick areas had larger values in images produced by the experimental system and smaller values in images from the GE system. The result of the segmenting software was a file for each subject, containing the coordinates of the lung contour each divided by the number of pixels in the $x$-dimension. The segmenting algorithm was implemented using the C++ language.

The lung contours for all the images were manually digitized from the background-subtracted, soft-tissue image. The digitizing software created files in the same format as the automatic software. A program read both contour files for each subject and then classified each pixel in the image in the four categories shown in Figure 6: true positive (TP), true negative (TP), false positive (FP), and false negative (FN), These were used to compute the accuracy, sensitivity, and specificity as follows:

$$Accuracy = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \quad (10)$$

$$Sensitivity = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (11)$$

$$Specificity = \frac{N_{TN}}{N_{TN} + N_{FP}} \quad (12)$$

The tests of the statistical shape model and reliability measure were performed using cross-validation with the complete data set as a partition.[19] The procedure is as follows:

1. Leave out one lung contour from the hand-digitized set.
2. Calculate the statistical shape model using the remaining contours in the hand-digitized set.
3. Compute the statistical shape model estimate (Eq 5) and the reliability measures (Eq 6 and scale parameter) with the automatically generated contour from the image of the case excluded in step 1.
4. Repeat steps 1 to 3 for each contour in the experimental set.

Note that the hand-digitized contours are used to generate the statistical shape model, but the output of the segmenting software is used to compute the model fit and the reliability measures.

## RESULTS

In this section, I describe the results of experimental tests of the segmenting method and the statistical reliability measure. The tests performed were as follows:

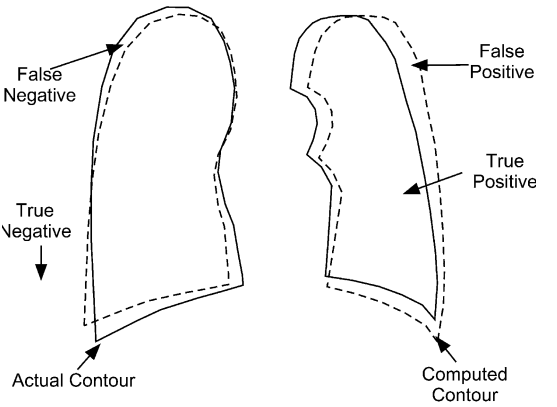


**Fig 6. Pixel classification. All pixels are classified into four categories: true positive (TP), true negative (TP), false positive (FP), and false negative (FN). The drawing shows the true contour (solid) and the computed contour (dashed) and examples of each class of pixels.**

1. Plots of the lung contours computed by the segmenting method superimposed on the soft-tissue images
2. Average value and standard deviation of the accuracy, sensitivity, and specificity
3. Plots of the first four modes of the statistical shape model
4. Distribution of the reliability measures and examples with individual images

### Lung Contours

Figure 7 shows the lung contours generated by the segmenting algorithm superimposed on soft-tissue images. The images from the GE system are at the top. Note the error in the image at the lower right corner. This was due to mis-positioning of the subject, placing the lung apex below the search region. The accuracy, sensitivity, and specificity of each image were computed using Eq 10, 11, 12. The average value and standard deviation of the performance measures were as follows:

Accuracy: 0.9747 ± 0.0101
Sensitivity: 0.9402 ± 0.0177
Specificity: 0.9875 ± 0.0107.

### Statistical Shape Model Results

The statistical shape model was computed using the hand-digitized contours from all the

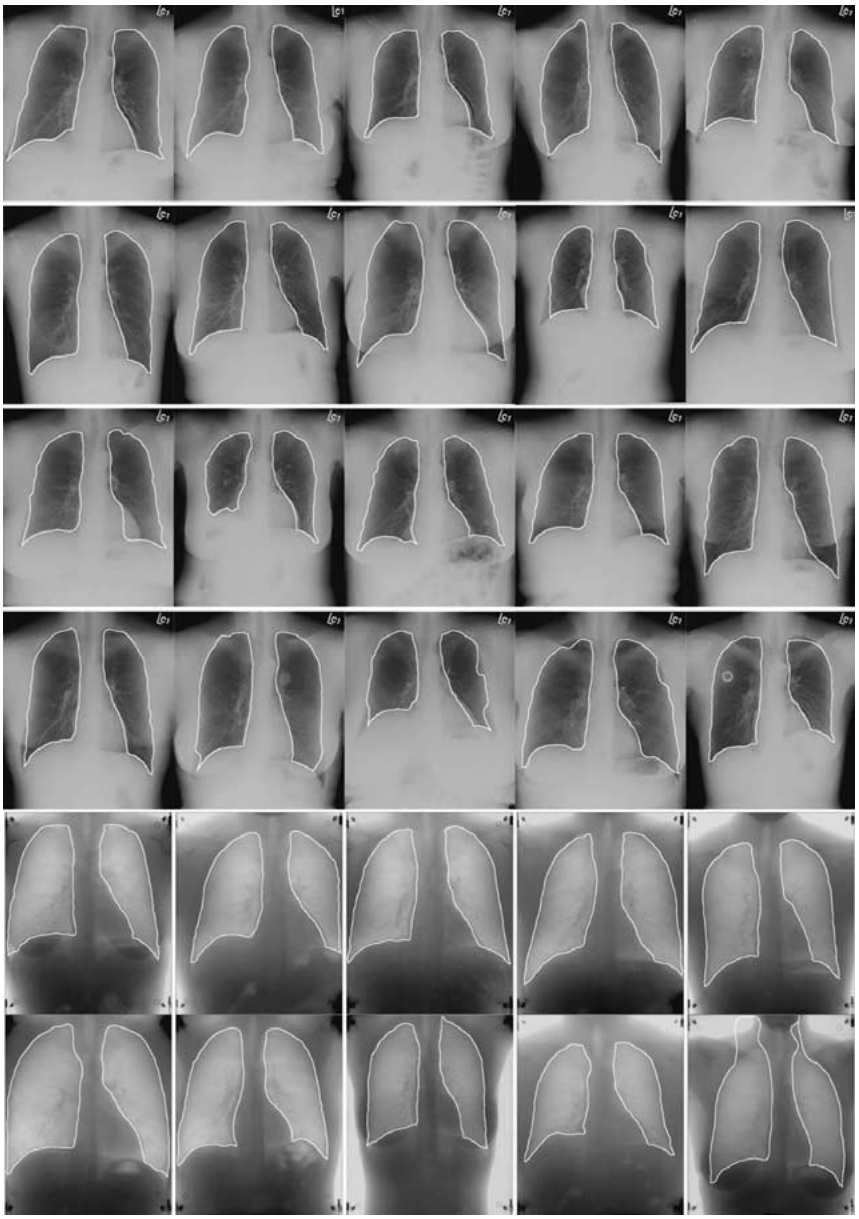

**Fig 7. Computed lung contours superimposed on the soft-tissue images. Note the processing error in the image in the lower right corner.**

images with Matlab software. Figure 8 shows plots of the eigenvalues (right axis) and the cumulative fraction of the total variance represented by the first $m$ modes (left axis). Over 96% of the total variance of the 30 lung contours can be represented by 10 modes. Figure 9 shows the variation of the lung shapes represented by the first four modes of the statistical shape model.

## Reliability Measure Results

Figure 10 shows a normal probability plot of values for the standardized residuals $S(Q)$ and the scale parameter $M$. This plot gives a graphical assessment of the accuracy of approximating the data with a normal distribution. If the data are normally distributed, the values will fall on a straight line. The plots show

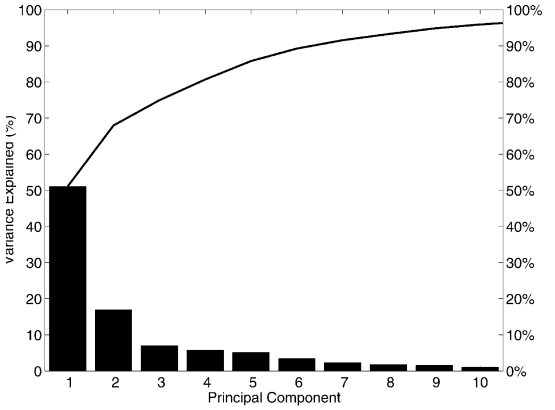


**Fig 8. Eigenvalues of the statistical shape model (left axis) and the cumulative fraction of the total variance represented by the first *m* modes (right axis). A total of 10 modes represents 96% of the total variance.**

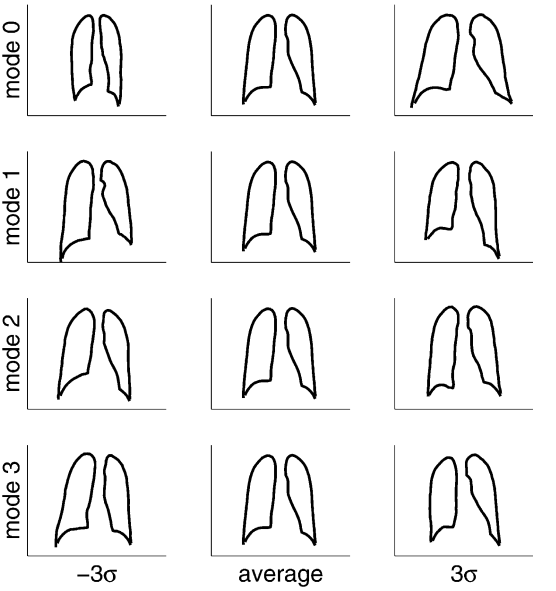


**Fig 9. Modes of the statistical shape model. Each of the four rows shows the contours represented by the average ± 3$\sigma$ (standard deviation) of a statistical shape model mode.**

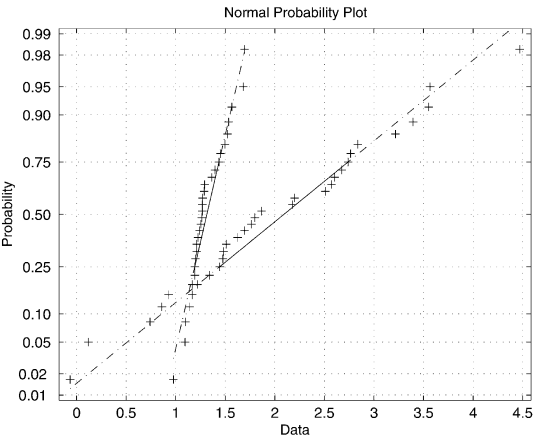


**Fig 10. Normal probability plots of the components of reliability measures. There are two superimposed plots. Each plot shows the experimental probability versus the data value for each point in the sample of a respective variable. The solid line is a fit to the 25th and 75th percentiles of the data and is therefore insensitive to the outliers. The *y*-scale is not uniform but is proportional to the probability that a normal random variable has that value; i.e., it is the inverse of the cumulative distribution function of a normal random variable.**

that normal distributions are good approximations for both variables.

The correlation coefficient of the two variables was 0.0190, indicating very low correlation. Because uncorrelated Gaussian random variables are also independent, the assumption in Eq 7 is justified.

By estimating the mean and standard deviation of normal approximations to the reliability measure distributions, Eq 7, 8, 9 can be used to compute the reliability measure $R$ of each lung contour. Figure 11 is a bar chart of the reliability measures on a logarithmic scale. Note that the first two values are substantially lower than the other values. Figure 12 shows the images for the first four cases in Figure 11, with the lung contours superimposed as a bright white line. The lowest two $R$ value images are on the left.

## DISCUSSION

Segmenting is the key initial step in CAD methods, and errors in this step could invalidate the overall results. Furthermore, CAD is often applied to screening examinations where the fraction of images with true positives is quite small, so a seemingly small error fraction is significant. Van Ginneken and Ter Haar Romeny[9] list a range of accuracies for segmenting methods reported in the literature. They also estimate the highest attainable accuracy of a segmenting method as the inter-observer variability. Table 1 compares the performance measures for the dual energy soft-tissue image segmenting method with the highest perform-

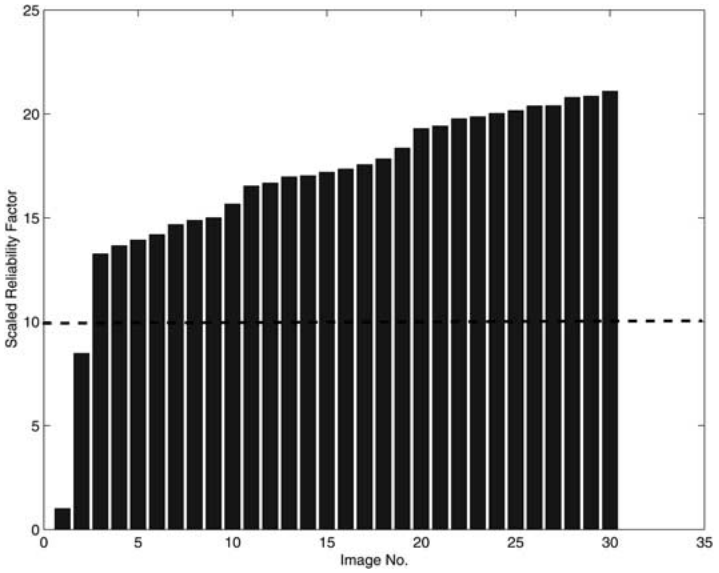


**Fig 11. Bar graph of the reliability measures for all the test images, sorted from smallest to largest. Note the step increase of the reliability measure after the first two images. The dashed line is an example of a threshold that could be defined so that values below the threshold are reported to the user. Two images fall below the threshold (see Fig 12).**

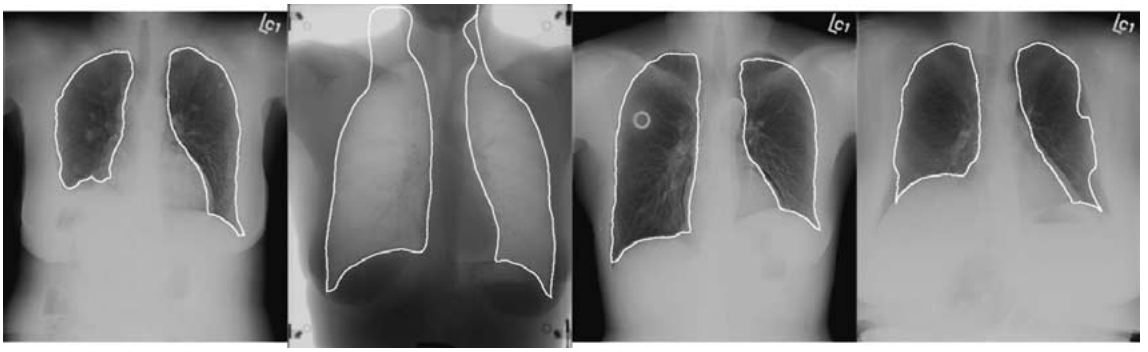

**Fig 12. The images for the first four cases in Figure 11. The computed lung outlines are the bright white lines. The two cases on the left have reliability measures that fall below the threshold, and the two on the right are above it. See the text for a discussion.**

ance method listed by van Ginneken and Ter Haar Romeny[9] the ''classification correction'' method. The dual energy method has higher accuracy (at the 5% confidence level) and specificity (at the 0.1% confidence level) using a one-sided *t*-test. The methods' sensitivities are equal at the$\alpha = 0.05$ level using a two-sided *t*-test.

As discussed by van Ginneken and Ter Haar Romeny,[9] improvements in performance should be considered in the context the fraction of the remaining errors eliminated compared to inter-observer variability, which is the best performance attainable. The dual-energy method described in this article eliminates 40% of the remaining errors as measured by accuracy and 70% of the errors as measured by specificity. These are large fractions of the remaining errors, and the dual-energy method makes a significant improvement in performance over existing methods.

Figure 11 shows that a threshold can be defined so that the reliability measure detects a patient with an unusual lung boundary con-

**Table 1. Segmenting Method Performance Measures**

| | Dual-Energy Soft-Tissue Image | Classification Correction[15] | Interobserver Variability[15] |
|---|---|---|---|
| Accuracy | 0.975 ± 0.0101 | 0.969 ± 0.00803 | 0.984 ± 0.00475 |
| Sensitivity | 0.940 ± 0.0177 | 0.943 ± 0.0330 | 0.957 ± 0.0174 |
| Specificity | 0.988 ± 0.0107 | 0.978 ± 0.0106 | 0.993 ± 0.00306 |

figuration or errors in the processing. The image with the lowest reliability measure (far left image in Figure 12) corresponds to a patient with an unusual right lung boundary and a high degree of asymmetry with the left lung. As discussed by Armato et al.[20] detection of cases with unusual anatomy has medical utility as an "automated triage" system. The reliability measure was also able to detect the one case among the 30 test images where the segmenting method had large errors, the second image from the left in Figure 12. The other two images on the right of Figure 12 had the next highest reliability measures, but they were above the threshold in Figure 11. In these cases, the segmenting method gave good results and the patients had normal lung configurations.

## CONCLUSIONS

The segmenting method described exploits the characteristics of dual-energy soft-tissue images to provide better performance results than previous methods. A reliability measure, using the standardized residuals and the scale parameter of a statistical shape model, detects unusual lung outlines and processing errors.

## ACKNOWLEDGMENTS

I acknowledge the assistance of Dr. Elizabeth H. Moore, Dr. J. Anthony Seibert of the University of California, Davis Medical Center, and Dr. Reginald Munden and Mr. Stephen K. Thompson of the University of Texas, M.D. Anderson Cancer Center in providing the human subject image data. I also acknowledge the assistance of Dr. Linda S. Maltz on statistical questions. This work was supported in part by National Institutes of Health SBIR grant R3CA97826A. Its contents are solely the responsibility of the author and do not necessarily represent the official views of the National Institutes of Health.

## REFERENCES

1. Nakamori, N, Doi, K, Sabeti, V, et al: Image feature analysis and computer-aided diagnosis in digital radiography: automated analysis of sizes of heart and lung in chest images. Med Phys 17:342-350, 1990

2. Sanada, S, Doi, K, MacMahon, H: Image feature analysis and computer-aided diagnosis in digital radiography: automated detection of pneumothorax in chest images. Med Phys 19:1153-1160, 1992

3. Ginneken, B van , Romeny, BMtH, Viergever, MA: Computer-aided diagnosis in chest radiography: a survey. IEEE TransMed Imaging 20:1228-1241, 2001

4. Lehmann, LA, Alvarez, RE, Macovski, A, Brody, WR, Pelc, NJ, Riederer, SJ, Hall, AL: Generalized image combinations in dual KVP digital radiography. Med Phys 8:659-667, 1981

5. Alvarez, RE, Macovski, A: Energy-selective reconstructions in X-ray computerized tomography. Phys Med Biol 21:733-744, 1976

6. Xu, X-W, MacMahon, H, Doi, K: Detection of lung nodule on digital energy subtracted soft-tissue and conventional chest images from a Cr system. In: Doi, K, MacMahon, H, Giger, M.L., Hoffmann, K.R. (eds) In Computer Aided Diagnosis in Medical Imaging Elsevier Science B.V., Amsterdam, The Netherlands, 1999, pp 63-69

7. Duryea, J, Boone, JM: A fully automated algorithm for the segmentation of lung fields on digital chest radiographic images. Med Phys 22:183-191, 1995

8. Vittitoe, NF, Vargas-Voracek, R, Floyd, CF Jr: Identification of lung regions in chest radiographs using Markov random field modeling. Med Phys 25:976-985, 1998

9. Ginneken, B van , terHaar Romeny, BM: Automatic segmentation of lung fields in chest radiographs. Med Phys 27:2445-2455, 2000

10. Cootes, TF, Taylor, CJ, Cooper, DH, Graham, J: Active shape models—their training and application. Computer Vision and Image Understanding 61:38-59, 1995

11. Ginneken, B: Computer-Aided Diagnosis in Chest Radiography. The Netherlands: Image Sciences Institute, University Medical Center Utrecht, 2001

12. Brown, MS, Wilson, LS, B.D., Doust, R.W., Gill, Sun, C: Knowledge based method for segmentation and analysis of lung boundaries in chest x-ray images. Comput Med Imag Graphics 22:463-477, 1998

13. Ginneken, B van , Frangi, AF, Staal, JJ, Romeny, BMtH, Viergever, MA: Active shape model segmentation with optimal features. IEEE Trans Med Imaging 21: 2002

14. Chotas, HG, Ravin, CE: Chest radiography: estimated lung volume and projected area obscured by the heart, mediastinum, and diaphragm. Radiology 193:403-440, 1994

15. Wuescher, DM, Boyer, KL: Robust contour decomposition using constant curvature criterion. IEEE Tr Pattern Analysis Mach Intell 13:41-51, 1991

16. Jackson, JE: A User's Guide to Principal Components. New York: John Wiley and Sons, 1991

17. Jackson, JE, Mudholkar, GS: Control procedures for residuals associated with principal component Analysis. Technometrics 21:341-349, 1979

18. Alvarez, RE, Seibert, JA, Poage, TF: Active dual-energy x-ray detector: experimental characterization. Proc SPIE 3032:419-426, 1997

19. Martinez, WL, Martinez, AR: Computational Statistics Handbook with MATLAB, Chapter 7. Boca Raton FL: CRC Press, 2001

20. Armato, SG, Giger, ML, MacMahon, H: Automated abnormal asymmetry detection in digital posteroanterior chest radiographs. In: Doi, K., MacMahon, H., Giger, M.L., Hoffmann, K.R. (eds) Computer-Aided Diagnosis in Medical Imaging. Elsevier Science B.V., Amsterdam, the Netherlands, 1999, pp 89-93