A novel multithreshold method for nodule detection in lung CT

Bruno Golosio,[a)] Giovanni Luca Masala, Alessio Piccioli,
Piernicola Oliva, and Massimo Carpinelli
Struttura Dipartimentale di Matematica e Fisica, Università di Sassari, via Vienna 2, 07100 Sassari
and INFN, Sezione di Cagliari, Cittadella Universitaria, 09042 Monserrato (CA), Italy

Rosella Cataldo
Dipartimento di Scienza dei Materiali, Università del Salento, 73100 Lecce, Italy and INFN,
Sezione di Lecce, Via Arnesano, 73100 Lecce, Italy

Piergiorgio Cerello
INFN, Sezione di Torino, Via Pietro Giuria 1, 10125 Torino, Italy

Francesco De Carlo
Dipartimento di Fisica, Università di Bari, via Amendola 173, 70126 Bari, Italy and INFN, Sezione di Bari,
Via Orabona 4, 70126 Bari, Italy

Fabio Falaschi
U. O. Radiodiagnostica 2, Ospedale Cisanello, Università di Pisa, via Paradisa 2, 56124 Pisa, Italy

Maria Evelina Fantacci
Dipartimento di Fisica, Università di Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy and INFN,
Sezione di Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy

Gianfranco Gargano
Dipartimento di Fisica, Università di Bari, via Amendola 173, 70126 Bari, Italy and INFN, Sezione di Bari,
Via Orabona 4, 70126 Bari, Italy

Parnian Kasae
Dipartimento di Fisica, Università di Siena, via Roma 56, 53100 Siena, Italy and INFN Sezione di Pisa,
Largo Pontecorvo 3, 56127 Pisa, Italy

Massimo Torsello
Unità Operativa di Radiologia, Ospedale Vito Fazzi, Piazza Muratore, 73100 Lecce, Italy

(Received 10 July 2008; revised 9 June 2009; accepted for publication 9 June 2009;
published 6 July 2009)

Multislice computed tomography (MSCT) is a valuable tool for lung cancer detection, thanks to its ability to identify noncalcified nodules of small size (from about 3 mm). Due to the large number of images generated by MSCT, there is much interest in developing computer-aided detection (CAD) systems that could assist radiologists in the lung nodule detection task. A complete multi-stage CAD system, including lung boundary segmentation, regions of interest (ROIs) selection, feature extraction, and false positive reduction is presented. The selection of ROIs is based on a multithreshold surface-triangulation approach. Surface triangulation is performed at different threshold values, varying from a minimum to a maximum value in a wide range. At a given threshold value, a ROI is defined as the volume inside a connected component of the triangulated isosurface. The evolution of a ROI as a function of the threshold can be represented by a treelike structure. A multithreshold ROI is defined as a path on this tree, which starts from a terminal ROI and ends on the root ROI. For each ROI, the volume, surface area, roundness, density, and moments of inertia are computed as functions of the threshold and used as input to a classification system based on artificial neural networks. The method is suitable to detect different types of nodules, including juxta-pleural nodules and nodules connected to blood vessels. A training set of 109 low-dose MSCT scans made available by the Pisa center of the Italung-CT trial and annotated by expert radiologists was used for the algorithm design and optimization. The system performance was tested on an independent set of 23 low-dose MSCT scans coming from the Pisa Italung-CT center and on 83 scans made available by the Lung Image Database Consortium (LIDC) annotated by four expert radiologists. On the Italung-CT test set, for nodules having a diameter greater than or equal to 3 mm, the system achieved 84% and 71% sensitivity at false positive/scan rates of 10 and 4, respectively. For nodules having a diameter greater than or equal to 4 mm, the sensitivities were 97% and 80% at false positive/scan rates of 10 and 4, respectively. On the LIDC data set, the system achieved a 79% sensitivity at a false positive/scan rate of 4 in the detection of nodules with a diameter greater than or equal to 3 mm that have been annotated by all four radiologists. [DOI: 10.1118/1.3160107]

Key words: computer-aided diagnosis (CAD), image processing, computed tomography (CT), image segmentation

I. INTRODUCTION

Lung cancer is the leading cause of cancer deaths. The 5-yr-survival rate is estimated to be only 16%.[1] The survival rate increases up to 49% for cases detected when the disease is still localized; however, only 16% of lung cancers are diagnosed at this early stage.[1] Screening tests for lung cancer focus on trying to detect the disease at an earlier and more curable stage, in particular, for high-risk individuals. In the past, chest radiography and sputum cytology were investigated as modalities for lung cancer screening,[2,3] showing a limited effectiveness in reducing lung cancer mortality. Low-dose helical computed tomography (CT) has provided promising results in the detection of early-stage lung cancer.[4–6] The large number of images that need to be interpreted by the radiologists in CT screening stimulated the development of various computer-aided detection (CAD) systems for lung nodules.[7–22] Some of the ongoing CT lung screening trials employ multislice CT (MSCT),[23,24] which, compared to single-slice helical CT, provides a higher resolution in the axial direction. The availability of tools for a truly three-dimensional data visualization and analysis is particularly important in order to take full advantage of the MSCT isotropic resolution. A complete CAD system for lung nodule detection, structured in four modules, is presented:

(1) Lung segmentation based on isosurface triangulation for modeling the pleural surface of the chest wall, and morphological closure operation for separating juxta-pleural nodules from the chest wall;
(2) regions of interest (ROI) selection based on isosurface triangulation with varying threshold (multithreshold method);
(3) computation of several features for each ROI as a function of the threshold; and
(4) ROI classification, based on self-organizing maps (SOMs) and artificial neural networks (ANNs).

II. DATA SETS

II.A. The Italung-CT data set

The Italung data set consists of 132 low-dose CT scans acquired by using a MSCT scanner (Somatom Plus 4 VZ, Siemens, Erlangen, Germany), according to a low-dose protocol at tube voltage of 140 kV, tube current of 20 mA, and slice thickness of 1.25 mm. The reconstruction was made at 1 mm interval using B50f reconstruction kernel. The scans were acquired at the Pisa center of the Italung-CT trial and annotated by expert radiologists.[24] The number of slices per scan ranges from 238 to 416. Each slice consists of a 512×512 pixel matrix, with pixel size ranging from 0.52 to 0.74 mm depending on the field of view.

Each scan of the data set was first annotated by one or two radiologists of the Pisa center of the Italung-CT trial. Each radiologist first performed a blinded review and identified the location of all abnormalities. For each suspect nodule, regardless of the size, the radiologist annotated the location and the diameter, as measured with electronic caliper. After the blinded review, 23 scans, separately annotated by two radiologists, were reviewed by the same two radiologists, who agreed on the nodules that were to be considered suspicious for cancer. The 23 scans were also reviewed by two experienced radiologists of the Ospedale Vito Fazzi (Lecce, Italy), who repeated the same procedure. The test set used to evaluate the performance of our CAD system includes all the nodules annotated by the radiologists of both institutions on the 23 scans in the unblinded reviews. All the nodules with diameter greater than or equal to 3 mm identified by one or two radiologists in the remaining 109 scans were used as training/optimization set. The numbers of nodules in the training set and in the test set were 176 and 45, respectively. The average nodule diameter and its standard deviation on the whole data set were 4.7 and 1.3 mm, respectively. The optimization of all the free parameters of the classification system was made on the training set, while the test set was only used for the evaluation of the system performance.

II.B. The LIDC data set

The CAD system was also tested using the scans provided by the Lung Image Database Consortium (LIDC) consortium,[25,26] a publicly available database of thoracic CT scans collected to promote the development of CAD systems and the comparison of their performances. Although our CAD system has been more specifically designed for thin slices CT scans, the test on the LIDC database is useful as a comparison with other systems on a public database. Furthermore, since each nodule larger than 3 mm in the LIDC database is described in detail by a set of radiologic characteristic, the performance of the CAD on different types of nodules can be studied.

Currently, 84 CT scans are available from the LIDC database, acquired using different types of CT scanners, single or multislice, with voltages between 120 and 140 kVp, x-ray-tube current between 40 and 422 mA, and slice thickness between 1.25 and 3 mm.

Each scan is provided together with its annotations by four experienced radiologists (each from a different institution), who identified the location and radiological characteristics of all the nodules and other anomalies larger than 3 mm in diameter, as well as the location of nodules smaller than 3 mm. Each radiologist performed first a blinded review of the scan, followed by an unblinded review with the information provided by the other radiologists available. No forced consensus was imposed. The annotations of all four radiologists in the unblinded review are available with the database. Only nodules larger than 3 mm were considered in this work.

III. METHOD

III.A. Isosurface triangulation

Given a density distribution $\mu(x,y,z)$, the isosurface corresponding to the value μ_I is defined as the set of points that satisfy the equation $\mu(x,y,z)=\mu_I$. Such equation defines a surface that separates volumes with density greater than μ_I from volumes with lower density. A discrete representation of the isosurface can be obtained by approximating it by a set of small triangular facets. Such a procedure is called *isosurface triangulation* and the representation obtained in this way is called *isosurface triangulated model*. Such models are widely used for volumetric data representation and analysis. One of the most popular algorithms for isosurface triangulation is the marching cube algorithm,[27,28] based on a division of the data volume in elementary cubes, followed by a standard triangulation inside each cube. In the original formulation, the marching cube algorithm makes use of 15 basic triangulations, and a total of 256 elementary triangulations is obtained from the basic ones by rotations, reflection, conjugation, and their combinations. The original formulation of the algorithm suffers from well-known problems of connectivity among triangles of adjacent cubes, solved in various ways.[29–31] We developed a variant of the marching cube algorithm that makes use of 21 basic triangulations, as shown in Fig. 1(a). Triangles of adjacent cubes are always well connected in this approach. Compared to other solutions, it is relatively simple and does not require substantial modifications of the original formulation of the algorithm; only the look-up table for the basic triangulations is changed.

The set of triangles yielded by the triangulation algorithm is divided into its connected components by forcing two triangles into the same component when they have at least one edge in common. In this way, the whole isosurface, which, in general is disconnected, is divided in a set of connected surfaces.

The surface-triangulation algorithm is complemented by an algorithm for the extraction of all the voxels inside each connected surface. In general, it is not trivial to write an efficient and robust algorithm for determining whether a point lies inside a closed triangulated surface.[32] In our approach, the volume inside each connected surface is extracted using a standard flood-fill algorithm,[33] starting from a seed voxel lying inside the surface, using a six-neighbor scheme and a simple bottom threshold inclusion rule (i.e., a neighboring voxel having a face in common with at least one voxel of the growing region is included in the growing region itself if its density is higher than μ_I). The basic triangulations in Fig. 1(a) are chosen in such a way that the set of voxels whose center lies inside a connected surface is the same set obtained by the flood-fill algorithm with the six-neighbor scheme. A set of tools for extracting statistical information on each connected surface and on the enclosed volume (surface area, volume, average and maximum density inside it, etc.) was implemented.

The robustness of the triangulation algorithm and its related tools was tested on the whole training set. As an example, Fig. 1(b) shows a triangulated model for the isosurface delimiting a juxta-pleural nodule. A skeleton representation was used in this case in order to outline the triangular mesh.

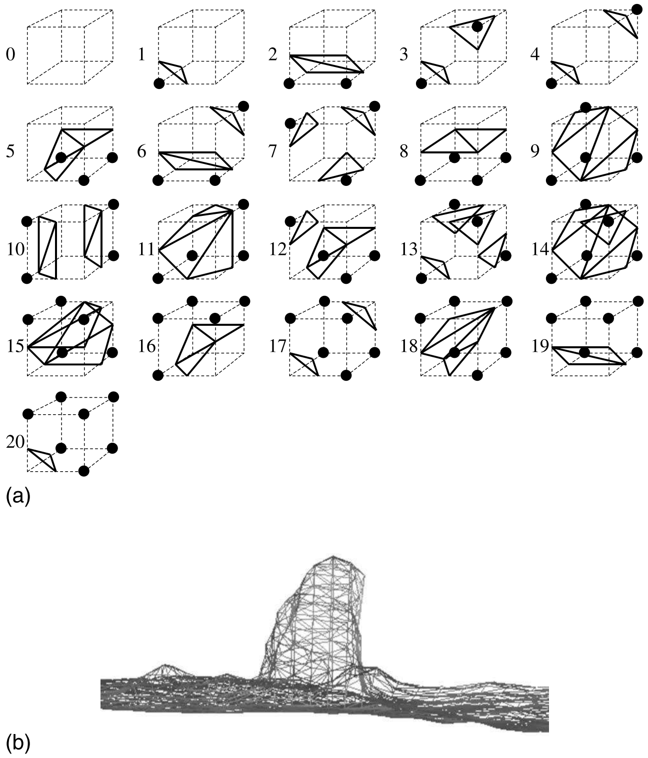


FIG. 1. (a) Basic combinations for the marching cube algorithm. Black circles represent voxels inside the isosurface. A standard triangulation of the surface is performed within each cube. All possible combinations are obtained from the fundamental ones by rotations and reflections. (b) Example of application of the isosurface-triangulation technique to a juxta-pleural nodule detected in a low-dose multislice lung CT scan. An isovalue of −400 in Hounsfield units has been used for the triangulation. A skeleton representation has been used in order to outline the triangle mesh.

III.B. Lung segmentation

The lung segmentation was performed using a method similar to those described in Refs. 34 and 35. The density of the lung parenchyma and that of the pleural surface of the chest wall are very different: about −700 and 0 in Hounsfield units (HU), respectively. The isovalue used to separate them, optimized on the training set, was $\mu_I=-500$ HU. Figure 2 shows an example of application of this technique. A mask for the lung parenchyma is obtained by flood filling the volume inside the isosurface. Since juxta-pleural nodules are connected to the pleural surface of the chest wall, they are not included in this mask. In order to identify them as ROIs, a morphological closure operation[36] is applied to the parenchyma mask, as shown in Fig. 3. An N-layer dilation is applied to the mask. The dilation also has the effect of smoothing the surface, removing high-curvature regions such as juxta-pleural nodules. The number of layers, i.e., the thickness of dilation, should be large compared to the nodule radius, and small compared to the radius of curvature of the pleural surface of the chest wall. The value $N=10$ in voxel units was empirically selected. An N-layer erosion is thus

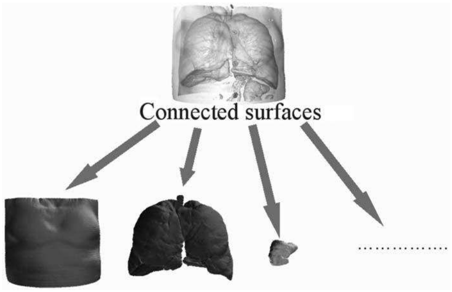


FIG. 2. Example of application of the isosurface-triangulation and connected component analysis to lung segmentation. An isovalue of −500 in Hounsfield units has been used for the triangulation. The surface separating the lung parenchyma from the chest wall is identified as the connected surface having the negative volume of largest magnitude.

applied, which basically brings the surface back to its original size, with high-curvature regions being removed. The eroded mask is merged with the original image by deleting all the voxels outside the mask. In this way, juxta-pleural nodules are included in the image obtained, while the chest wall is excluded. This simple approach works well far from the hilar region of the lungs and from the region where the two lungs are connected to each other. Lung segmentation can be improved by taking special care for such regions.[37]

III.C. ROI selection

The region of interest (ROI) selection procedure consists in identifying regions in the parenchyma mask that can contain suspect lung nodules. The isosurface-triangulation method is generally suitable to accomplish this task. Using this approach, the ROIs are represented by the collection of the connected surfaces corresponding to a given isovalue or, equivalently, by the connected volumes inside them.

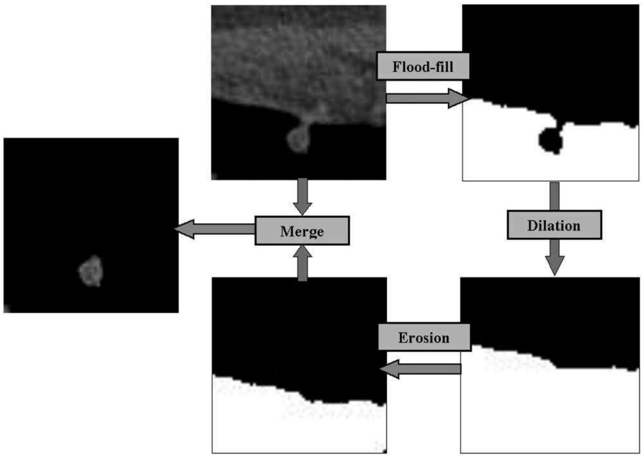


FIG. 3. Morphological closure operation used to separate juxta-pleural nodules from the pleural surface of the chest wall: A mask for the lung parenchyma is obtained by flood filling the volume inside the isosurface separating the lung parenchyma from the chest wall; a 10-layer dilation is applied to the mask, removing high-curvature regions such as juxta-pleural nodules; a 10-layer erosion is thus applied, bringing the surface back to its original size, with high-curvature regions being removed; the eroded mask is merged with the original image by deleting all voxels outside the mask. In this way, juxta-pleural nodules are included in the image obtained, while the chest wall is excluded.

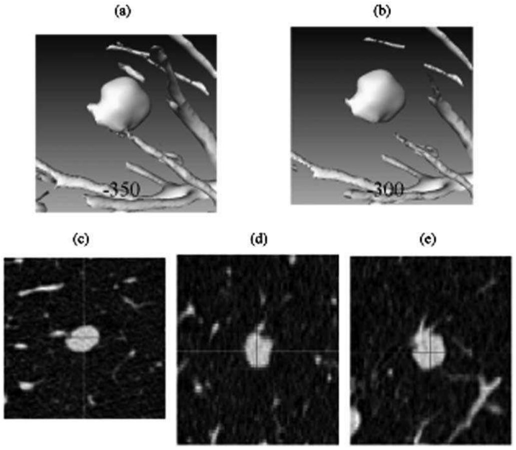


FIG. 4. [(a) and (b)] Three-dimensional views of isosurfaces corresponding to a nodule connected to a blood vessel for thresholds of −350 and −300 in Hounsfield units, respectively. [(c)–(e)] Axial, sagittal, and coronal view of the same nodule.

Since the nodule density is higher than that of lung parenchyma, the surface representing their border can be reconstructed by setting a proper intermediate isovalue. Internal, isolated nodules are easily identified by this simple criterion since they correspond to connected volumes above the threshold. Juxta-pleural nodules will also correspond to connected volumes above the threshold, as long as the lung segmentation works well. However, the lung segmentation sometimes leaves part of the pleura inside the volume of interest, especially in high-convexity regions. Juxta-pleural nodules in these regions will remain connected to part of the pleura. In addition, regardless of the segmentation performance, filtering effects imply that a layer along the pleural surface of the chest wall stays inside the volume of interest, as its density is higher than the parenchyma average density. If the threshold is too low, juxta-pleural nodules will be connected to this layer. Furthermore, internal nodules are not always isolated, as they can establish connections with the blood vessels. Figures 4 and 5 show some examples of nodules connected to the blood vessels, segmented using different threshold values. If the threshold is too low, these nodules appear as connected to the vessels, showing up as a single big ROI. On the other hand, the threshold must not be too high. If it becomes higher than the density of a nodule, part of this nodule will be lost and its volume will be underestimated. In the extreme case when the threshold is higher than the maximum density inside the nodule, it will be completely lost. The threshold selection is always a trade-off between the need to separate nodules from other structures and the requirement of not loosing part of the nodule itself. Tests on real cases show that there is no single threshold suitable to all nodules. Possible solutions to the problem of threshold optimization could be based on adaptive thresholds. However, such approaches must deal with the problem that the value of some features (volume, surface, roundness, etc.) critically depends on the threshold itself. A multithreshold

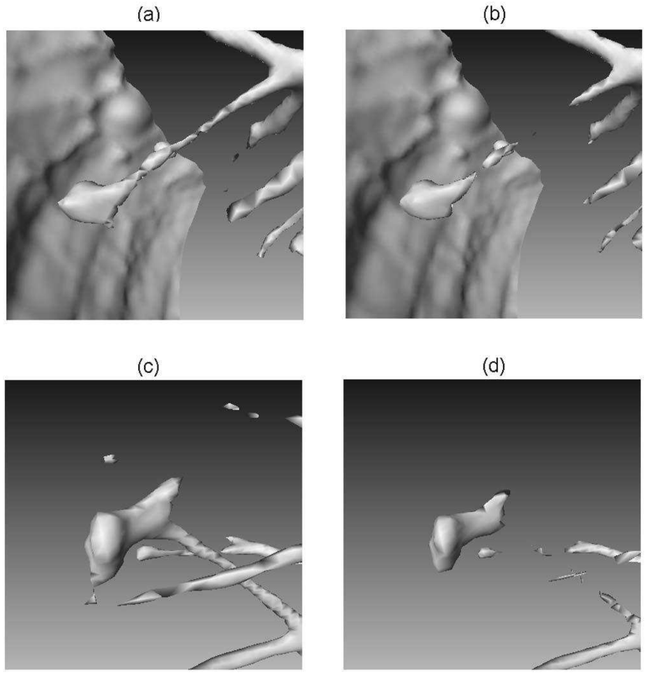


FIG. 5. Three-dimensional views of the isosurfaces corresponding to two nodules connected to a blood vessel, one segmented at the thresholds of −350 and −300 in Hounsfield units [(a) and (b), respectively], the other segmented at the thresholds of −250 and −150 in the same units [(c) and (d), respectively].

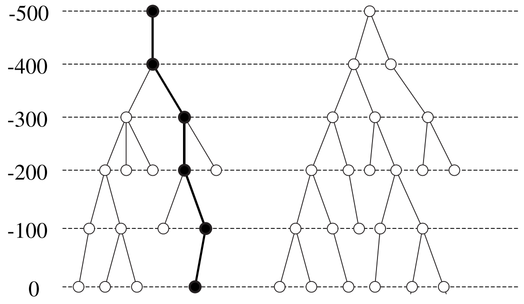


FIG. 6. Schematic example of the treelike structure of the correspondence between ROIs in the multithreshold analysis. The circles represent the ROIs, while the segments represent correspondence between ROIs at neighboring threshold. Threshold values are in Hounsfield units. A MROI, such as the one represented by full circles, is defined as a path on one of these trees, which starts from a terminal ROI and ends up on the root ROI.

approach for the automated detection of lung nodules in CT scans was first proposed by Armato *et al.*[38] The authors applied 36 Gy level thresholds to the segmented lung volume. For each threshold, they identified contiguous structures with associated gray levels greater than the threshold and observed that single structures identified at lower gray-level threshold value can disassociate into multiple smaller structures at higher threshold values. A structure is identified as a nodule candidate if its volume is smaller than a predefined value. For each nodule candidate, a feature vector is calculated for the ROI corresponding to a single threshold value, the one in which the volume criterion was first satisfied. In this sense this method is similar to adaptive threshold approaches.

The solution explored in this work is a complete multithreshold analysis. First, the images are filtered by a three-dimensional Gaussian filter, with $\sigma=1$ in pixel units. This value is kept low so as to avoid the suppression of small nodules. Although the filter removes details of the nodule border shape, it has the positive effect of reducing the signal in the connections between the nodule and other structures, such as the vessels and the chest wall, while keeping the signal high in the nodule core. After this preprocessing stage, the threshold is varied from a minimum to a maximum value in a wide range, and the ROI selection is performed for each value in this range. In order to follow the evolution of a ROI when the threshold increases, we need to match ROIs corresponding to neighboring thresholds. Each ROI at a given threshold is completely included in some ROI at the neighboring lower threshold. However, this correspondence is not one-to-one. There are different possible cases:

(1) Only one ROI at the higher threshold corresponds to the ROI at the lower threshold;
(2) the ROI at the lower threshold splits itself as the threshold grows, so the corresponding ROIs at the higher threshold are more than 1; and
(3) the higher threshold is greater than the maximum density inside the ROI at the lower threshold, therefore this ROI will have no correspondence in the higher threshold; in this case the ROI is said to be a terminal ROI.

Therefore the connections among ROIs at neighboring thresholds have a treelike structure (see Fig. 6). A multithreshold ROI (MROI) is defined as a path on this tree, which starts from a terminal ROI and ends on the root ROI (see Fig. 6). A MROI maps the evolution of a ROI as the threshold varies from the minimum to the maximum value. In the treelike representation, the MROI corresponding to a nodule connected to a blood vessel and the MROI corresponding to the vessel itself have the root ROI in common, i.e., they belong to the same tree. However, in the case of solid nodules, at some threshold value the ROI corresponding to the nodule and the one corresponding to the vessel will separate from each other, yielding two different branches of the tree. Therefore, the nodule and the vessel correspond to two different MROIs. For example, for the three nodules shown in Figs. 4 and 5, branching occurs between −350 and −300 HU [Figs. 4(a) and 4(b)], between −350 and −300 HU [Figs. 5(a) and 5(b)], and between −250 and −150 HU [Figs. 5(c) and 5(d)], respectively. A single threshold lower than −350 HU would not detect the three nodules: They would not be separated from the vessels. On the other hand, using a single threshold higher than −350 HU many low-density nodules would be lost. Using a multithreshold method, solid nodules connected to the blood vessels as well as low-density nodules can be detected. The assumption that there is a threshold that allows separation of the nodules from vessels, however, is a limitation of the proposed method: While this is generally true in the case of solid nodules, the assumption may not hold for subsolid nodules.

As described in Sec. III D, several features are computed for each MROI as functions of the threshold. Some of them, such as volume, surface, and roundness, critically depend on the threshold. The MROI representation provides much more information than the ROI representation. The feature functions associated with blood vessels and with different types of nodules show different behaviors. The task of learning how to distinguish nodules from other structures via the behavior of the feature functions is assigned to the classification system.

III.D. Feature extraction

Some ROI features are used as input to the classification system: volume, roundness, maximum density, mass, and principal moments of inertia. The volume V and the surface S are computed via the triangulated model of the ROI surface. More specifically, the surface is defined as the sum of the surfaces of all triangles of the triangulated model; the volume is computed using the method described in Ref. 39. The R_V (R_S) radius of a sphere with the same volume (surface) of the ROI are defined as follows:

$$R_V = \left(\frac{3V}{4\pi}\right)^{1/3}, \quad R_S = \sqrt{\frac{S}{4\pi}}.$$

The ROI roundness is defined as the ratio

$$\text{Roundness} = \frac{R_V}{R_S},$$

which is a real number in the $[0,1]$ range. This ratio is the square root of conventional sphericity ($\pi^{1/3}(6V)^{2/3}/S$). In order to avoid errors due to discretization in roundness calculation, the volume is computed on the triangulated model (using an algorithm for computing the volume inside a polyhedron) rather than simply counting the number of voxels inside the ROI surface and multiplying it by the voxel volume. Such errors can be relatively large, particularly when the ROI contains a small number of voxels.

The mass, the center of mass, and the inertia tensor are computed using standard formulas. The mass and the center of mass are defined as

$$M = \sum\nolimits_{k=1}^{N} m_k, \quad R_i = \frac{\sum\nolimits_{k=1}^{N} m_k r_{k,i}}{M},$$

where the index k runs over all the voxels inside the ROI surface, m_k is the density associated with the voxel k (in HU) times the voxel volume (in mm^3), i is equal to 1, 2, and 3 for x, y, and z, respectively, and $r_{k,i}$ is the position vector of the voxel k center (in mm). The inertia tensor is defined as

$$I_{i,j} = \sum_{k=1}^{N} m_k (r_k'^2 \delta_{ij} - r'_{k,i} r'_{k,j})$$

with $r'_{k,i} = r_{k,i} - R_i$ being the position vector of the voxel k center with respect to the center of mass. The moment of inertia tensor has been diagonalized so as to find the three principal moments of inertia I_1, I_2, and I_3. The maximum density is defined as the maximum value of the density associated with all voxels inside the ROI surface (in HU).

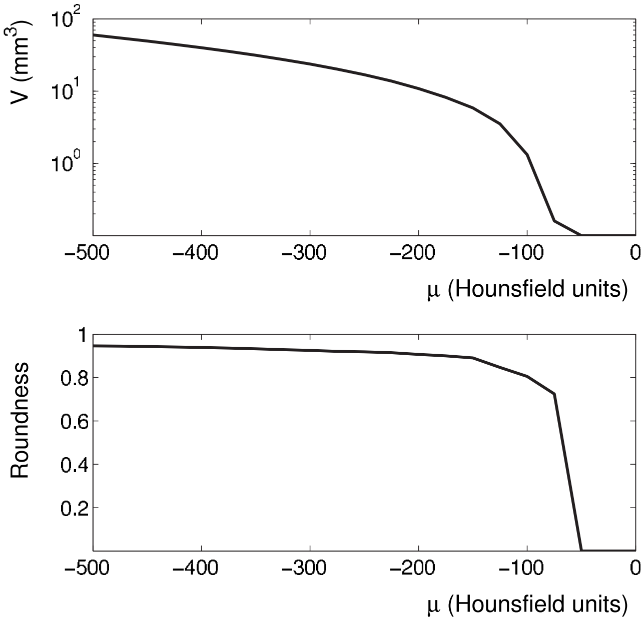


FIG. 7. Volume and roundness of a MROI corresponding to an internal, isolated nodule, plotted as functions of the threshold. When the threshold approaches the maximum intensity of the ROI, the volume decreases abruptly to zero. Roundness has a high value in the whole range.

Lung nodules are considered to be significant if their diameter is at least 3 mm. Their shape is roughly spherical, therefore their roundness must be close to 1 and their three principal moments of inertia should be similar. Many ROIs obtained from the ROI selection procedure are due to noise in the image. However, their volume and density are generally lower than those of the actual nodules. Other ROIs correspond to blood vessels. Their volume can be very large, but their roundness is generally much lower than that of the nodules, and the principal moment of inertia corresponding to the axis parallel to the vessel is generally much smaller than the other two. In the MROI approach, each feature is a function of the threshold. The volume of MROIs corresponding to isolated nodules shows a simple behavior (see Fig. 7): It starts from a value of the order of 10^2 mm^3 for the lower threshold and decreases slowly as the threshold increases, until the threshold approaches the maximum intensity, and the volume abruptly decreases to zero. The roundness value is high in the whole range. The behavior of the same functions for an internal nodule connected to the blood vessels is peculiar, as shown in Fig. 8. For a low threshold the nodule appears as connected to the vessels and is part of a single, very big ROI with a volume of about 10^5 mm^3. Its roundness, being mostly influenced by the blood vessels contribution, is quite low. As the threshold increases, the nodule separates from the vessels at a certain critical value and the ROI volume abruptly decreases to a value of the order of 10^2 mm^3. At the same time, the roundness will rapidly increase. Figure 4 shows a three-dimensional view of the iso-

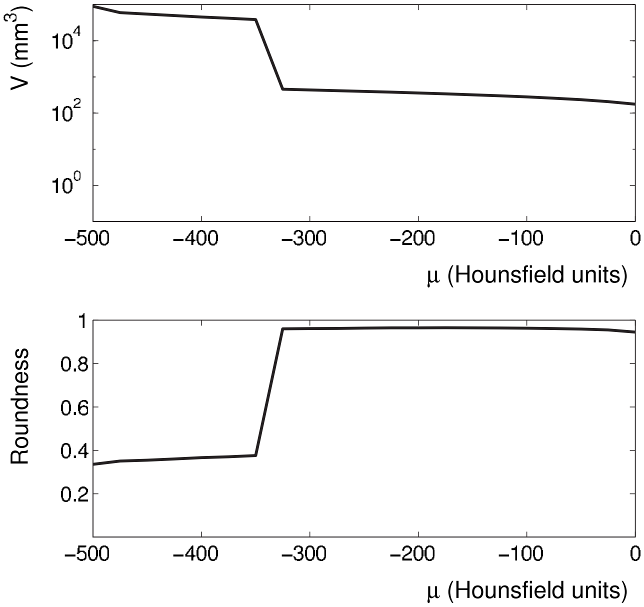


FIG. 8. Volume and roundness of a MROI corresponding to a nodule connected to the blood vessels, plotted as functions of the threshold. For a low threshold nodule and vessels are segmented as a single ROI having large volume and low roundness. Above some critical threshold, the nodule is separated from the vessels; therefore, the volume decreases abruptly to a lower value and the roundness increases abruptly to a higher value.

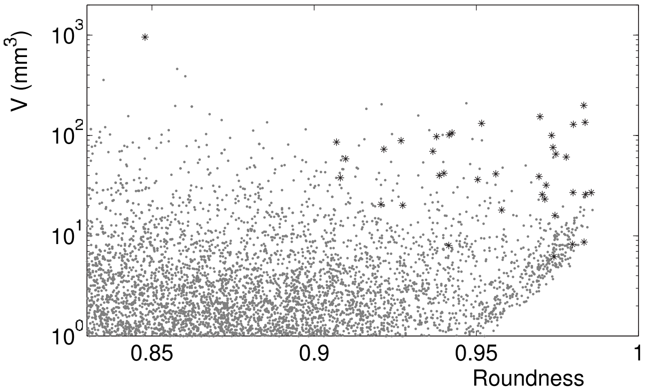
surface corresponding to the same nodule, for thresholds of −350 and −300 HU, together with axial, sagittal and coronal view of the nodule.

Figures 9(a) and 9(b) show, for all the test set ROIs, the scatterplots of the volume versus the roundness for thresholds of −400 and 0 HU, respectively. The dots represent false positives (FPs), while the stars represent the ROIs corresponding to lung nodules. Some nodules are connected to the blood vessels at the lower threshold value (see also Fig. 4), thus they do not appear in Fig. 9(a). On the other hand, the higher threshold is greater than the maximum density of some nodules, which do not appear in Fig. 9(b).

III.E. Classification

The MROI features as functions of the threshold can be used as input to a classification system. The size of the input vector is equal to the number of features times the number of sampling steps in the analyzed threshold range. For this study, six sampling points, from −500 to 0 HU with a sampling step of 100 HU, were used. If a ROI disappears at a threshold smaller than or equal to 0 HU, the corresponding elements of the feature vector are set to zero.

(a) Threshold = –400



(b) Threshold = 0

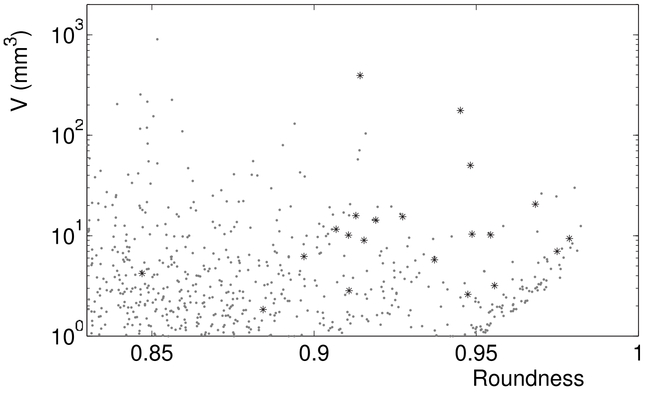


FIG. 9. [(a) and (b)] Scatterplot of the ROI volume versus roundness evaluated for a threshold of −400 and 0 in the Hounsfield units, respectively. The dots represent false positives, while the stars represent the ROIs corresponding to lung nodules.

The classification system is based on ANNs. The number of MROIs obtained by the ROI selection procedure is large: 46 024 in the whole training set. Since the training-validation procedure used by the classification system does not work properly if the number of negative examples is much greater than the number of positive examples, it is important to downsample the negative examples without altering their distribution in the feature space. For such purpose a two-dimensional SOM[40,41] was used. The SOM divides the examples in representative classes, each output node of the SOM corresponding to a class. The number of classes cannot be too small, since all types of false positives should be represented by some class; neither it can be too large since each class should contain a statistically significant number of examples. Considering the size of our database, a compromise was found to be 10 × 10 output nodes, which yields an average of 460 negative examples per output node. The set of negative examples is downsampled to 1060 cases by taking about 2% of the cases from each representative class, i. e. from each output node of the SOM. The positive examples in the training set are 176.

The classification stage is done using a feedforward neural network with one hidden layer and logistic activation function. The number of input neurons is 43 (seven features, each one evaluated for six threshold values, plus the maximum density inside the MROI). Two output neurons are used, corresponding to a positive or to a negative response. The network was trained through the standard back-propagation learning algorithm with momentum.[41] A *leave-one-out* cross validation procedure[40,42] was used on the training set for the evaluation of the root mean square error, which was used for the optimization of the number of learning epochs and of the number of neurons in the hidden layer.

TABLE I. Total number of nodules in the Italung-CT test set with diameter greater than or equal to 3 and 4 mm and ROI selection procedure sensitivity. The lower and upper limits of 95% confidence intervals are evaluated through a method described by Wilson (Ref. 43) with correction for continuity.

Nodule diameter (mm)	Total No. of nodules	True positives	ROI selection sensitivity
≥3	45	42	$0.93^{+0.05}_{-0.13}$
≥4	30	30	$1.00_{-0.14}$

TABLE II. CAD system sensitivity evaluated on the Italung-CT test set for nodules with diameter greater than or equal to 3 and 4 mm at false positive rates of 10 and 4 per scan.

Nodule size (mm)	FP/CT	True positives	Sensitivity
≥3	10	38/45	$0.84^{+0.09}_{-0.14}$
	4	32/45	$0.71^{+0.12}_{-0.16}$
≥4	10	29/30	$0.97^{+0.03}_{-0.16}$
	4	24/30	$0.80^{+0.12}_{-0.19}$

The optimal number of hidden neurons was found to be 11.

III.F. Large nodules

While all nodules in the Italung-CT data set are relatively small, the largest diameter being 9.85 mm, some of the nodules in the LIDC data set are not. In particular, for 16 nodules the diameter is larger than 20 mm. Although generally the recognition of the larger nodules is relatively easy for the classification system, their segmentation requires special care, particularly when they are connected to the chest wall. The number of layers used in the morphological closure procedure, optimized for small nodules, can be insufficient for larger nodules, which may not be correctly segmented. For this reason, big nodules were treated separately. The segmentation procedure is the same as described for the smaller nodules, except for the following:

(a) The images are filtered by a two-dimensional Gaussian filter for suppressing the signal in the connections between nodules and other structures, while keeping it high in the solid part of the nodules. The sigma value of the filter, optimized for the big nodules, was 4 in pixel units.

(b) The optimal number of layers used in the two-dimensional morphological closure operation for the detection of big nodules was 40.

(c) The use of an ANN classifier was not appropriate for big nodules because the number of positive examples, i.e., the number of nodules larger than 20 mm, is too low. Therefore, the candidates were selected with two simple thresholds on the ROI volume and roundness.

The whole CAD system is thus composed by two subsystems, one specialized in the detection of nodules larger than 20 mm, the other specialized in the detection of smaller nodules. The global false positive rate is the sum of the rates of the two subsystems.

IV. RESULTS

IV.A. Results on the Italung-CT data set

The ROI selection procedure applied to the 23 CTs of the test set yielded a total number of 12 043 ROIs, 42 of them being true positives, 12 001 false positives. Since the number of nodules with diameter greater than or equal to 3 mm located by the radiologists in the test set is 45, the sensitivity of the ROI selection procedure is 93%. The diameter of the three nodules that are not detected by the ROI selection procedure is smaller than 4 mm. The average number of FPs per CT at this stage is 522. Table I shows the sensitivity of the ROI selection procedure on the test set nodules with diameter greater than or equal to 3 and 4 mm.

The ROIs detected by the selection procedure are used as input to the ANN classifier. Figure 10 shows the FROC curves for the test set nodules with diameter greater than or equal to 3 and 4 mm. The sensitivities of working points, selected on the FROC curves, are reported on Table II.

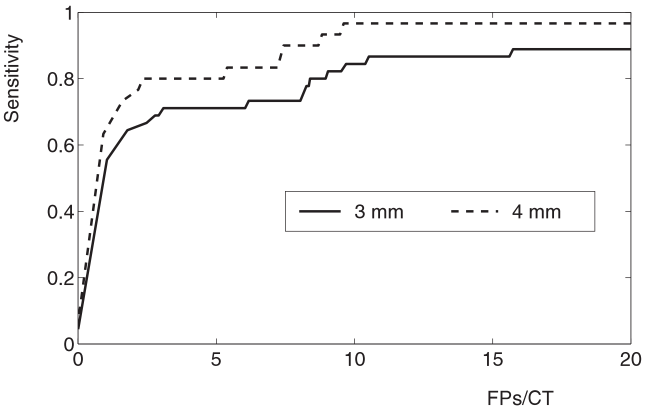


FIG. 10. FROC curves of the CAD system for the nodules of the Italung-CT test set having diameter greater than or equal to 3 mm (solid line) and 4 mm (dashed line).

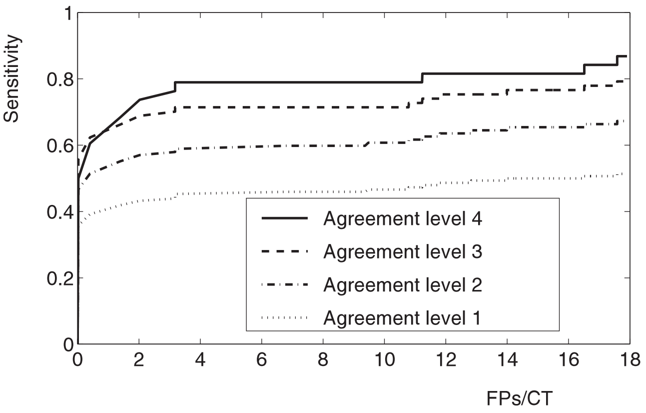


FIG. 11. FROC curves of the CAD system for the nodules of the LIDC data set with different agreement levels and diameter greater than or equal to 3 mm.

TABLE III. CAD system sensitivity on the nodules of the LIDC data set with diameter of ≥3 mm obtained at a false positive rate of 4 per scan at different agreement levels.

	True positives	Sensitivity
Agreement level 1	67/148	$0.45^{+0.08}_{-0.08}$
Agreement level 2	63/107	$0.59^{+0.09}_{-0.10}$
Agreement level 3	55/77	$0.71^{+0.09}_{-0.12}$
Agreement level 4	30/38	$0.79^{+0.11}_{-0.17}$

IV.B. Results on the LIDC data set

The CT scan parameters and the procedure used to construct the reference standard for the LIDC data set are different from the ones used for the Italung-CT database. Therefore, a new optimization and training procedure of the classification system were necessary. One of the differences that mostly affected the CAD performance was the slice thickness. Since the slice thickness of the scans in the LIDC database is not constant, it was used as an additional input feature for the classification system.

After finding the intersections among the nodules identified by different radiologists, the nodules larger than 3 mm were divided in four subsets: 41/30/40/37 nodules annotated by one/two/three/four out of four radiologists, respectively. These numbers show a substantial inter-reader variability, reflecting a difference in opinions even among experienced radiologists. Each of these subsets has been divided randomly in two subgroups containing almost the same number of nodules, one used as training set, the other as test set. Subsequently, the role of the training and test sets was exchanged and the whole procedure was repeated. In this way, the CAD system was tested on all the nodules of the data set. There are different ways to build a reference standard from the annotations of the four radiologists. We used the method suggested by Opfer *et al.*[21] and called *ground truth with agreement level j* the list of all the nodules marked by at least *j* of the 4 radiologists.

Figure 11 shows the FROC curves of our CAD system for the nodules with different agreement levels and diameter greater than or equal to 3 mm. Table III shows the CAD system sensitivities on the four ground truth lists, obtained at a false positive rate of 4 per scan. Clearly, the sensitivity is higher for nodules identified with higher consensus among the four radiologists.

V. DISCUSSION

Figure 12 shows the CAD system sensitivity on the LIDC data set, obtained at a false positive rate of 4 per scan and averaged over the four radiologists readings, as a function of some of the nodule characteristics. A strong correlation between the CAD sensitivity and the radiologist perception of subtlety, i.e., difficulty of detection, can be noticed from the histogram in Fig. 12(a). Figure 12(b) shows that the CAD sensitivity for nodules that are suspicious for malignancy is higher than for nodules that are unlikely to be malignant. Figure 12(c) shows that the CAD performance on solid nodules is higher than on part-solid and nonsolid nodules.

Table IV compares the results of other CAD systems to the one described in this work. An objective comparison is difficult because the minimum diameter of the nodules, the number of slices per CT, the size of the databases, and other CT scan parameters are different. References 15 and 19 provide sensitivity and number of false positives per CT similar to the ones that we obtain for nodules larger than 4 mm. However, it should be observed that the number of scans in the database used in Ref. 19 is too small to divide it in a training/validation set and an independent test set.

Sahiner *et al.*[20] reported the performance of two algorithms, a region growing method and an active contour model, on a test set including 33 scans from the University of Michigan and 29 scans from the LIDC database. Their analysis only includes internal nodules having a diameter of ≥3 mm that were not ground-glass opacities. The combination of the two algorithms shows a sensitivity of 70% for nodules with agreement level 1 at a false positive rate of 1.5 per scan. However, since juxta-pleural nodules and ground-glass opacities have been excluded from this analysis, the results cannot be easily compared to the ones we obtain.

Opfer *et al.*[21] presented a CAD system with a validation study on the LIDC data set and discuss how the performance of their algorithm is influenced by the choice of the underlying ground truth. For nodules having a diameter of ≥4 mm that have been simultaneously confirmed by all four radiologists (agreement level 4) their CAD system shows a sensitivity of 91% at a false positive rate of 4 per scan. The superior performance of their system may be related to the use of a special filter, based on distance transformation, which identifies all structures similar to circles or half circles, and to the use of highly discriminating features.

From the histogram in Fig. 12(c), it appears that the performance of our system would improve if we could increase the detection rate of nonsolid and part-solid nodules. The use of a more specialized filter in the preprocessing step and of additional features, such as the ones used by Opfer *et al.*, would probably improve the system sensitivity. On the other hand, most of the false positives detected by our system are due to imperfections of the lung segmentation in regions where the pleural surface of the chest wall is irregular or have a high curvature. Therefore, in order to reduce the false positive rate without decreasing the sensitivity, the lung segmentation procedure should be improved.

VI. CONCLUSION

A complete CAD system for lung nodule detection has been developed and its performance evaluated with a data set of low-dose multislice CT scans. The advantage of a multithreshold surface-triangulation approach was explored using a treelike structure to follow the evolution of a ROI at varying threshold, introducing the concept of MROI and building the input to the classification system by computing features as sampled functions of the threshold value.

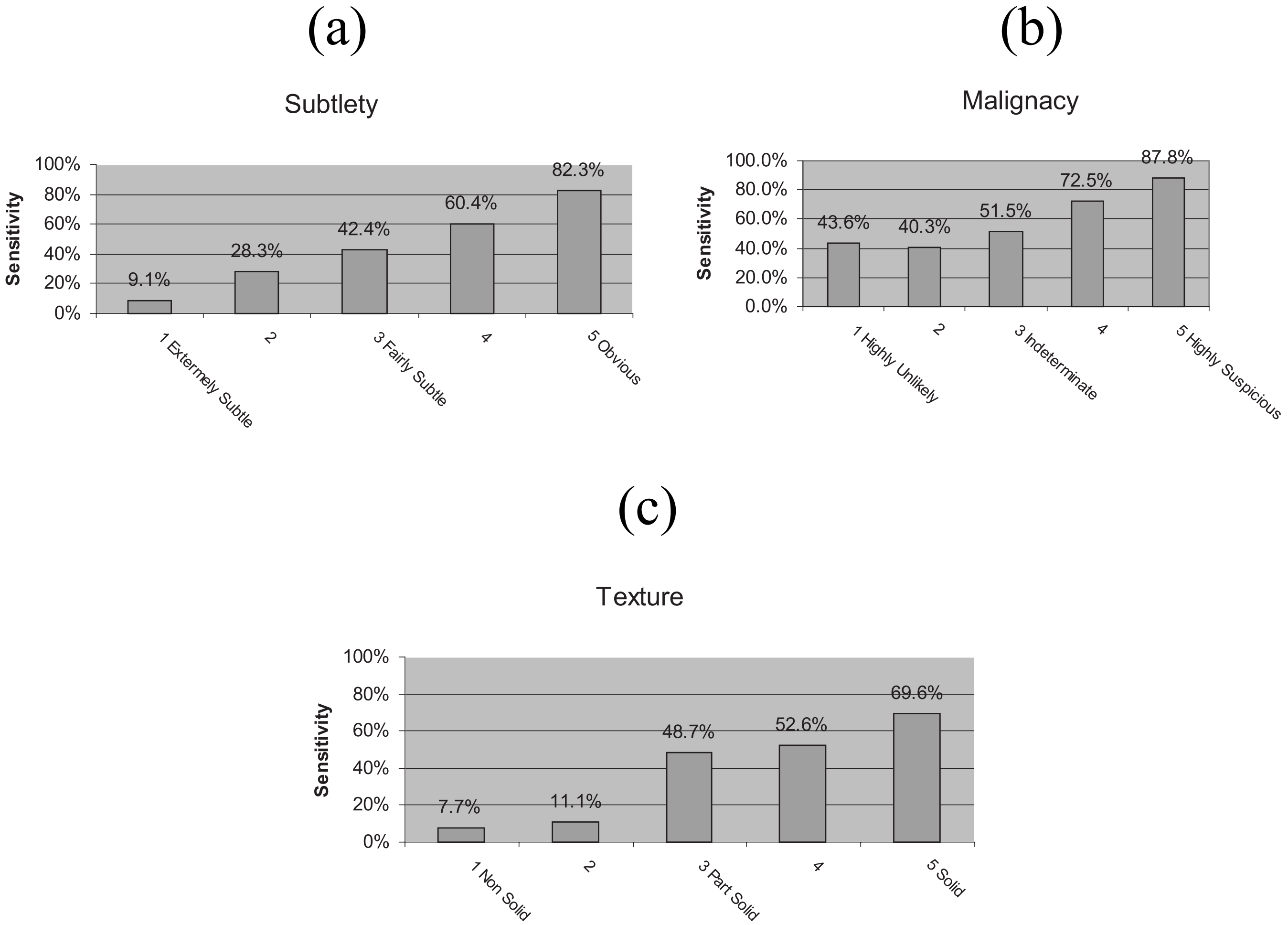


FIG. 12. Sensitivity of the CAD system evaluated on the LIDC data set at a false positive rate of 4 per scan as a function of some of the nodule characteristics. Each distribution is averaged over the four radiologists readings. (a) Sensitivity as a function of nodule subtlety, i.e., difficulty of detection. (b) Sensitivity as a function of nodule malignancy. (c) Sensitivity as a function of nodule texture.

The system achieved sensitivities of 84% and 71% at false positive rates of 10 and 4 per scan, respectively, for nodules with a diameter greater than or equal to 3 mm on the Italung-CT test set, and a 79% sensitivity at 4 false positives per scan on the LIDC data set for nodules with a diameter greater than or equal to 3 mm that were annotated by all four radiologists. In our opinion, other nodule segmentation methods, different from surface triangulation, could benefit from using a similar multithreshold approach by using other critical parameters instead of the surface-triangulation threshold, for instance, the parameter used for the halting criterion in some active contour models or the threshold on the merit function used for the inclusion rule in some region growing methods.

ACKNOWLEDGMENTS

The work presented in this paper has been carried out in the framework of the INFN (Istituto Nazionale di Fisica

TABLE IV. Results reported by different CAD systems and by the one described in this work.

CAD	Scans (train+test)	Diameter (mm)	Sensitivity (%)	FPs/CT
Reference 15	38+63	≥4	80.3	4.8
Reference 16	38	≥3	70	14.4
Reference 19	15	≥5	88.5	6.6
Reference 20	94+62	≥3	70	1.5
Reference 21	91	≥4	91	4
Our CAD (Italung)	109+23	≥3	71	4
		≥4	80	4
Our CAD (LIDC, agreement level 4)	84	≥3	79	4

Nucleare)-funded MAGIC-5 (Medical Applications in a GRID Infrastructure Connection) project. Some of the results reported in this paper have been produced in the framework of the Cybersar project, managed from the COSMOLAB Consortium and partially funded from the Ministero dell'Università e della Ricerca (MIUR) by the Piano Operativo Nazionale "Ricerca Scientifica, Sviluppo Tecnologico, Alta Formazione" (PON 2000–2006) (http://www.cybersar.com). The authors would like to thank Gavino Cherchi for helping with the installation and configuration of the computer cluster. They also thank Dr. A. De Liperi, Dr. C. Spinelli of the U.O. Radiodiagnostica 2, Azienda Ospedaliera Universitaria Pisana, Dr. M. Barattini, Professor D. Caramella, Dr. T. Tarantino of the Dipartimento di Radiologia Diagnostica ed Interventistica, Università di Pisa, and Dr. I. Zecca, Dr. A. Margiotta, Dr. M. Foti, and Dr. S. Cazzetta of the Unità Operativa di Radiologia Ospedale Vito Fazzi, Lecce.

[a] Electronic mail: golosio@uniss.it

[1] Cancer Facts & Figures, American Cancer Society, 2007.

[2] R. S. Fontana *et al.*, "Screening for lung cancer: A critique of the Mayo lung project," Cancer **67**, 1155–1164 (1991).

[3] J. K. Frost *et al.*, "Early lung cancer detection: Results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study," Am. Rev. Respir. Dis. **130**, 549–554 (1984).

[4] M. Kaneko, K. Eguchi, H. Ohmatsu, R. Kakinuma, T. Naruke, K. Suemazu, and N. Moriyama, "Peripheral lung cancer: Screening and detection with low-dose spiral CT versus radiography," Radiology **201**, 798–802 (1996).

[5] S. Sone, F. Li, Z. G. Yang, S. Takashima, Y. Maruyama, M. Hasegawa, J. C. Wang, S. Kawakami, and T. Honda, "Characteristics of small lung cancers invisible on conventional chest radiography and detected by population based screening using spiral CT," Br. J. Radiol. **73**, 137–145 (2000).

[6] S. Diederich *et al.*, "Screening for early lung cancer with low-dose spiral CT: Prevalence in 817 asymptomatic smokers," Radiology **222**, 773–781 (2002).

[7] M. N. Gurcan *et al.*, "Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system," Med. Phys. **29**, 2552–2558 (2002).

[8] S. G. Armato III, A. Roy, H. Macmahon, F. Li, K. Doi, S. Li, and M. B. Altman, "Evaluation of automated lung nodule detection on low-dose computed tomography scans from a lung cancer screening program," Acad. Radiol. **12**, 337–346 (2005).

[9] S. G. Armato III *et al.*, "Automated lung nodule classification following automated nodule detection on CT: A serial approach," Med. Phys. **30**, 1188–1197 (2003).

[10] S. G. Armato III, M. B. Altman, and P. J. LaRivière, "Automated detection of lung nodules in CT scans: Effect of image reconstruction algorithm," Med. Phys. **30**, 461–472 (2003).

[11] H. Arimura, S. Katsuragawa, K. Suzuki, J. Shiraishi, S. Li, and K. Doi, "Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening," Acad. Radiol. **11**, 617–629 (2004).

[12] J. S. Kim, J. H. Kim, G. Cho, and K. T. Bae, "Automated detection of pulmonary nodules on CT images: Effect of section thickness and reconstruction interval—initial results," Radiology **236**, 295–299 (2005).

[13] K. T. Bae, J. S. Kim, Y. H. Na, K. G. Kim, and J. H. Kim, "Pulmonary nodules: Automated detection on CT images with morphologic matching algorithm—preliminary results," Radiology **236**, 286–294 (2005).

[14] M. S. Brown, J. G. Goldin, S. Rogers, H. J. Kim, R. D. Suh, M. F. McNitt-Gray, S. K. Shah, D. Truong, K. Brown, J. W. Sayre, D. W. Gjertson, P. Batra, and D. R. Aberle, "Computer-aided lung nodule detection in CT: Results of large-scale observer test," Acad. Radiol. **12**, 681–686 (2005).

[15] K. Suzuki, S. G. Armato III, F. Li, S. Sone, and K. Doi, "Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography," Med. Phys. **30**, 1602–1617 (2003).

[16] A. S. Roy, S. G. Armato III, A. Wilson, and K. Drukker, "Automated detection of lung nodules in CT scans: False positives reduction with the radial-gradient index," Med. Phys. **33**, 1133–1140 (2006).

[17] T. W. Way, L. M. Hadjiiski, B. Sahiner, H. P. Chan, P. N. Cascade, E. A. Kazerooni, N. Bogot, and C. Zhou, "Computer-aided diagnosis of pulmonary nodules on CT scans: Segmentation and classification using 3D active contours," Med. Phys. **33**, 2323–2337 (2006).

[18] O. Osman, S. Ozekes, and O. N. Ucan, "Lung nodule diagnosis using 3D template matching," Comput. Biol. Med. **37**, 1167–1172 (2007).

[19] R. Bellotti *et al.*, "A CAD system for nodule detection in low-dose lung CTs based on region growing and a new active contour model," Med. Phys. **34**, 4901–4910 (2007).

[20] B. Sahiner, L. M. Hadjiiski, H. P. Chan, J. Shi, T. Way, P. N. Cascade, E. A. Kazerooni, C. Zhou, and J. Wei, "The effect of nodule segmentation on the accuracy of computerized lung nodule detection on CT scans: Comparison on a data set annotated by multiple radiologists," Proc. SPIE **6514**, 65140L.1–65140L.7 (2007).

[21] R. Opfer and R. Wiemker, "Performance analysis for computer aided lung nodule detection on LIDC data," Proc. SPIE **6515**, 65151C.1–65151C.9 (2007).

[22] C. S. White, R. Pugatch, T. Koonce, S. W. Rust, and E. Dharaiya, "Lung nodule CAD software as a second reader: A multicenter study," Acad. Radiol. **15**, 326–333 (2008).

[23] C. A. van Iersel, H. J. de Koning, G. Draisma, W. P. T. M. Mali, E. T. Scholten, K. Nackaerts, M. Prokop, J. D. F. Habbema, M. Oudkerk, and R. J. van Klaveren, "Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON)," Int. J. Cancer **120**, 868–874 (2007).

[24] G. Picozzi, E. Paci, A. Lopez Pegna, M. Bartolucci, G. Roselli, A. De Francisci, S. Gabrielli, A. Masi, N. Villari, and M. Mascalchi, "Screening of lung cancer with low dose spiral CT: Results of a three year pilot study and design of the randomised controlled trial 'Italung-CT'," Radiol. Med. (Torino) **109**, 17–26 (2005).

[25] S. G. Armato III *et al.*, "Lung image database consortium: Developing a resource for the medical imaging research community," Radiology **232**, 739–748 (2004).

[26] M. F. McNitt-Gray *et al.*, "The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation," Acad. Radiol. **14**, 1464–1474 (2007).

[27] W. E. Lorensen and H. E. Cline, "Marching Cubes: A high resolution 3D surface construction algorithm," Comput. Graph. **21**, 163–169 (1987).

[28] B. Golosio, A. Brunetti, R. Cesareo, S. R. Amendolia, D. V. Rao, and S. M. Seltzer, "Images of soft materials: A 3D visualization of interior of the sample in terms of attenuation coefficient," Nucl. Instr. Methods Res. Phys. A **465**, 577–583 (2001).

[29] A. Gueziec and R. Hummel, "Exploiting triangulated surface extraction using tetrahedral decomposition," IEEE Trans. Vis. Comput. Graph. **1**, 328–342 (1995).

[30] K. S. Delibasis, G. K. Matsopoulos, N. A. Mouravliansky, and K. S. Nikita, "A novel and efficient implementation of the marching cubes algorithm," Comput. Med. Imaging Graph. **25**, 343–352 (2001).

[31] X. Renbo, L. Weijun, and W. Yuechao, "A robust and topological correct marching cube algorithm without look-up table," *Proceedings of the Fifth International Conference on Computer and Information Technology* (IEEE, Shanghai, 2005), p. 565–569.

[32] J. Lane, B. Magedson, and M. Rarick, "An efficient point in polyhedron algorithm," Comput. Vis. Graph. Image Process. **26**, 118–125 (1984).

[33] P. Heckbert, *A Seed Fill Algorithm, Graphics Gems* (Academic, Boston, 1990), pp. 275–277.

[34] S. Krass, D. Selle, D. Boehm, H. H. Jend, A. Kriete, and H. O. Peitgen, "Determination of bronchopulmonary segments based on HRCT data," *CARS* (Elsevier, Amsterdam, 2000), pp. 584–589.

[35] T. Tozaki, Y. Kawata, N. Niki, H. Ohmatsu, K. Eguchi, and N. Moriyama, "Three-dimensional analysis of lung areas using thin slice ct images," in *Physiology and Function from Multidimensional Images*, edited by E. A. Hoffman (SPIE, Medical Imaging, Newport Beach, CA, 1996), pp. 2–11.

[36] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing* (Prentice-Hall, Englewood, Cliffs, NJ, 1988).

[37] G. De Nunzio, E. Tommasi, A. Agrusti, R. Cataldo, I. De Mitri, M. Favetta, R. Bellotti, S. Tangaro, N. Camarlinghi, and P. Cerello, "An

innovative lung segmentation algorithm in CT images with accurate delimitation of the hilus pulmonis," *IEEE Nuclear Science Symposium Conference Record 2008* (IEEE, Dresden, Germany, 2008), pp. 5359–5361.

[38] S. G. Armato III, M. L. Giger, and H. MacMahon, "Automated detection of lung nodules in CT scans: Preliminary results," Med. Phys. **28**, 1552–1561 (2001).

[39] J. O'Rourke, *Computational Geometry in C*, 2nd ed. (Cambridge University Press, Cambridge, 1998), p. 26.

[40] O. Duda, P. E. Hart, and D. G. Stark, *Pattern Classification*, 2nd ed. (Wiley, New York, 2001).

[41] S. Haykin, *Neural Networks—A Comprehensive Foundation*, 2nd ed. (Prentice-Hall, Englewood, Cliffs, NJ, 1999).

[42] S. R. Amendolia, A. Brunetti, P. Carta, G. Cossu, M. L. Ganadu, B. Golosio, G. M. Mura, and M. G. Pirastru, "A real time classification system of thalassemic pathologies based on artificial neural networks," Med. Decis Making **22**, 18–26 (2002).

[43] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," J. Am. Stat. Assoc. **22**, 209–212 (1927).