

# Rapid image recognition of body parts scanned in computed tomography datasets

Volker Dicken · B. Lindow · L. Bornemann ·  
J. Drexl · A. Nikoubashman · H.-O. Peitgen

Received: 10 January 2008 / Accepted: 17 December 2009 / Published online: 30 May 2010  
© CARS 2010

## Abstract

**Aim** Automatic CT dataset classification is important to efficiently create reliable database annotations, especially when large collections of scans must be analyzed.

**Method** An automated segmentation and labeling algorithm was developed based on a fast patient segmentation and extraction of statistical density class features from the CT data. The method also delivers classifications of image noise level and patient size. The approach is based on image information only and uses an approximate patient contour detection and statistical features of the density distribution. These are obtained from a slice-wise analysis of the areas filled by various materials related to certain density classes and the spatial spread of each class. The resulting families of curves are subsequently classified using rules derived from knowledge about features of the human anatomy.

**Results** The method was successfully applied to more than 5,000 CT datasets. Evaluation was performed via expert visual inspection of screenshots showing classification results and detected characteristic positions along the main body axis. Accuracy per body region was very satisfactory in the trunk (lung/liver >99.5% detection rate, presence of abdomen >97% or pelvis >95.8%) improvements are required for zoomed scans.

**Conclusion** The method performed very reliably. A test on 1,860 CT datasets collected from an oncological trial showed

that the method is feasible, efficient, and is promising as an automated tool for image post-processing.

**Keywords** Classification · Computed tomography · Anatomy · Body region · Database · Segmentation · Content-based image retrieval

## Introduction

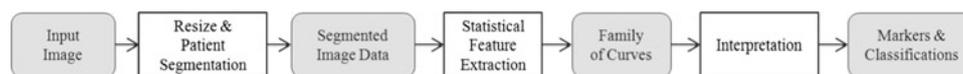
Our method was developed to address the problem of classifying a very large collection of some thousand CT image datasets with respect to recognition of the depicted body region. This was motivated by lack or inconsistency of the available documentation in cases collected over many years in various research projects with international partners to allow efficient search for test and evaluation cases required in ongoing research projects. The aim is to rely as much as possible on image data and to largely avoid the use of DICOM tag information found to be unreliable, inconsistent, or unavailable due to anonymization in many cases. It turned out that DICOM tag information is hard to classify for a collection of data originating from multiple countries where different languages were used for entering DICOM tags. This is especially true for contrast use, often recorded if at all in conjunction with a large variety of abbreviations. Also body region classification of DICOM tags based on DICOM tag values faces difficulties, as the respective DICOM tag “BODY REGION” is often not correctly set and scan protocol names such as “biph. Thorax^Abdomen” give no reliable clue whether chest or abdomen is actually depicted. Frequently, certain universal protocols are clinically used to scan body regions not directly related to the protocol name. We aim for a coarse automatic classification on a level that a human with basic knowledge of CT

---

JCARS paper submission to CARS 2008, 2nd revision of 27 October 2009.

---

V. Dicken (✉) · B. Lindow · L. Bornemann · J. Drexl ·  
A. Nikoubashman · H.-O. Peitgen  
Fraunhofer MEVIS, Institute for Medical Image Computing,  
Bremen, Germany  
e-mail: volker.dicken@mevis.fraunhofer.de  
URL: <http://www.mevis.fraunhofer.de/>



**Fig. 1** Principal data flow of the classification method

**Table 1** Density ranges and corresponding color codes in the subsequent figures and curves

Material class	Density range	Color code
Background Voxels below $-1,000$ HU after areas outside of the patient were set to below $-1,000$ HU	Below $-1,000$ HU	Grey
Air or other gases Mostly within the body after patient segmentation	Up to $-910$ HU	Violet
Lung parenchyma (or foam in positioning aids if not removed)	Up to $-300$ HU	Dark blue
Fatty tissue	Up to $-50$ HU	Dark green
Water dense areas	Up to $+30$ HU	Brown
Not enhanced soft tissue	Up to $+60$ HU	Orange
Enhanced soft tissue	Up to $+100$ HU	Yellow
Contrasted vessels (may include strongly contrasted liver and partial volume voxels within or near bones)	Up to $+250$ HU	Pink
Highly contrasted areas and most bones	Up to $500$ HU	Light blue
Harder bone parts and concentrated digestive contrast	Up to $1,200$ HU	Light green
Solid parts of large bones, skull and metal implants	Above $1,200$ HU	White

Another class used is the collection of all dense materials (every voxel  $> -300$  HU, curves in olive). The peak value of this along the scan length provides a reliable estimate of the patient size for images containing torso cross-sections

imaging could easily perform within seconds. However, we require a fast and robust method that allows processing of a CT scan of some 100 MB in a few seconds once images are loaded. Stability and careful memory management are required to run the procedure overnight on large image collections comprising of some Terabyte of collected data. Our aim is somewhat similar to the object of the content-based image retrieval works by Greenspan [1,2] or Lehmann [3,4] and others cited there working on image database retrieval. Though, unlike these publications on 2D imaging we are dealing with often very large 3D datasets and require very fast determination of a coarse level classification. The closest related work appears to be by a Japanese group at Gifu University [5].

## Materials and Methods

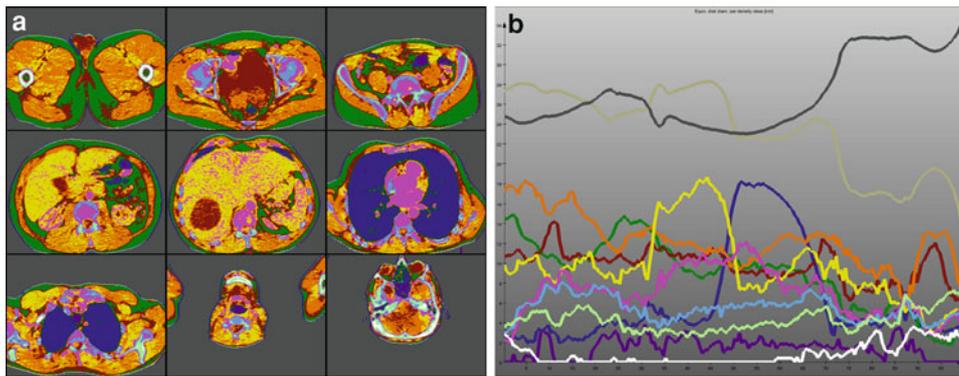
The method comprises a fast slice-wise patient segmentation that attempts to remove most of the background (air, table, clothing, positioning aids) followed by a slice-wise extraction of statistical features for a range of density classes on which the classification is based (Fig. 1).

### Extraction of statistical density class features

The statistical feature extraction provides for selected slices (e.g. every 3 mm) estimates of the area filled with certain materials for a number of density classes (e.g., lung parenchyma, fat, soft tissue, bones, c.f. Table 1). Further, the spatial spread of each class is computed per slice as mean position (center of gravity) and standard deviation of the  $X$  and  $Y$  coordinates of those voxels classified to belong into the corresponding density class. The starting point of the body region classification are families of curves showing the effective diameter of the area filled per sampled slice (i.e., diameter of an area equivalent disk) as well as the spread parameters for each investigated material class plotted over the table position ( $Z$ -axis). In order to differentiate various levels of contrasted tissue and normal soft tissue, a fine classification of individual voxels into 11 materials was employed (Fig. 2).

### Noise level estimation and noise removal

To reduce the influence of image noise, especially in lung or bone CT reconstructions, and to estimate the noise level a local average smoothing was employed prior to the



**Fig. 2** **a** Classified materials at nine selected body cross sections from the CT component of a PET/CT scan. **b** Family of curves showing the effective diameter (0–34 cm) of the area filled by each density class over the table position (0–100 cm; for the color codes per material class c.f. Table 1)

classification of individual voxels. For each material class a noise level was computed from the mean of the absolute difference between original and locally smoothed voxel-values. The mean was computed separately for all materials over all voxels with  $3 \times 3$  kernel averaging values in the respective density class. We selected the noise level of the non-enhanced soft tissue fraction (e.g., muscles, heart, brain, liver without contrast) as basis for a classification of the noise level, since many other material classes fill only a small area or include a high fraction of partial volume with high gradients that yield high differences after smoothing, and some classes are not present in every region of the body. The noise classification distinguishes data with (a) little image noise, (b) normal noise as in most CTs scanned with normal radiation dose and reconstructed for high soft tissue contrast, (c) high noise as is typical for lung and bone reconstructions, and (d) very high noise, almost exclusively present in low-dose scans of the thorax.

#### Patient segmentation

Our algorithm to determine the body region from the statistical density class feature curves will start with a search for lung tissue within the scan. In order to avoid erroneously taking CT-table, clothing, blankets, positioning aids or partial volume voxels near the skin for lung tissue, it is necessary to remove most of the background prior to the feature extraction. This patient segmentation step is implemented on a per-slice basis to efficiently integrate with the statistical feature extraction also working on a per-slice basis for selected positions only. Hence, an arbitrarily large dataset may be analyzed keeping only one slice at a time in memory. The main idea of the patient segmentation consists of a column-wise search for the patient contour, defined as the first voxel exceeding a minimum tissue density ( $-600$  HU from the front,  $-300$  HU from the back) followed by a minimum length (4 mm from the chest, 10 mm from the back) of voxels

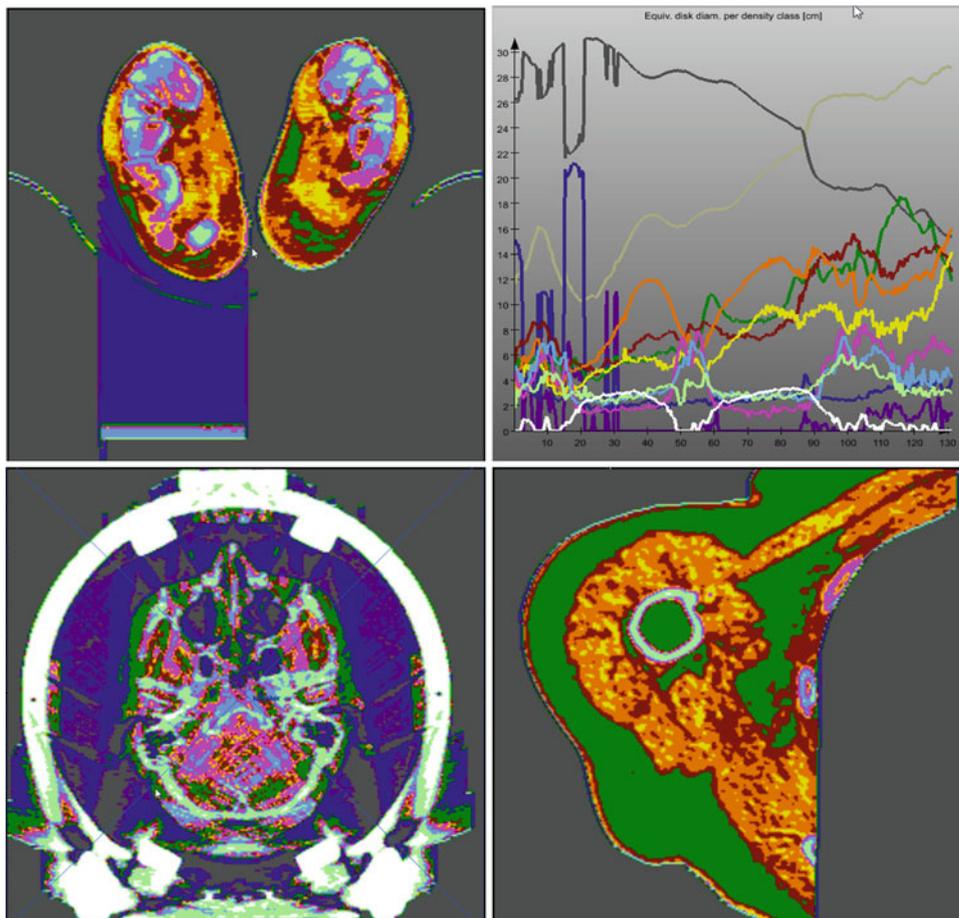
exceeding the threshold density. The search will usually run over clothing or thin tubes on the patient's front and the plastic shell of the CT's patient table. The different treatment of front and back was found to be advantageous to at least partially keep thin structures like ears and nose, while being able to remove the table and thicker blankets at the back. Some heuristic modifications were added to the basic idea to avoid major jumps in the contour due to mistaking the edge of a thicker amount of soft tissue dense material within the table (e.g. for holding positioning aids attached to the table) for the patient contour. Furthermore, a special case is required to avoid removing lung tissue near the image boundary in zoomed images of the heart because it was confused with foam in the table or positioning aids. For proper segmentation of the patient also in noisy low dose data an inexpensive cumulative line-by-line smoothing was integrated with the segmentation:

$$\begin{aligned} \text{Smoothed value} &:= \\ &(\text{original value} + \text{previous voxels smoothed value})/2 \end{aligned} \quad (1)$$

Patient segmentation may fail, e.g., in scans with circular reconstruction area when the reconstruction is zoomed on the lungs or if thicker dense material, e.g., stereotaxie equipment is present (c.f. Fig. 3).

#### Performance issues

In order to speed up the calculation it was empirically determined that the curve families providing the basis of the subsequent classification will hardly change if only a small subset of the available data is processed. It turned out that it suffices to smooth and classify, e.g., only every third voxel in every third row of one slice every 3 mm of data initially down-sampled to  $256^2$  resolution using nearest neighbor sampling. Due to smoothing with a  $3 \times 3$  kernel averaging every voxel of the



**Fig. 3** Errors resulting from limitations of our patient contouring or zoomed views. *Top row* pseudo lung near feet resulting from failed complete removal of a support cushion on top of a dense thick board, a configuration somewhat similar to chest wall; Classified voxels in feet assumed to be patient (*left*), resulting eff. diameter curves of area filled by different material classes (*right*, for colors codes c.f. Table 1,

*dark blue lung, grey background*). *Bottom row* pseudo lung resulting from mistaking a stereotaxic device for the patient and imaging artifacts increasing the air density in the intermediate space (*left*), erroneously removed lung, mistaken as foam in table in zoomed view of a shoulder (*right*)

$256^2$  intermediate images will still contribute to the result. For the actual classification of the smoothed voxel values a pre-computed look-up table was employed. Thus, computing the initial statistical feature curves from the CT data may be performed within times below 1 s even for CT data with several hundred images exceeding 200 MB, as computational effort only depends on the length of the scan range, and not on the actual number of images in the CT series. With these down-sampling and local smoothing settings, the computation is mainly limited by the hard disc or network speed that restricts the number of volume dataset slices that may be analyzed within a given time.

#### Classification of body regions

The classification employs a complex set of rules with parameters selected heuristically, and derived either from knowledge about the human anatomy or selected after experiments

on limited selections of initially failed classifications. Core decisions follow the pseudo code

- 
1. **Start:** Analyze *Lung* related curves
  2. **IF** (*no Lung* found) **OR** (*upper Lung* found)  
    **THEN** Analyze *Head* related curve features
  3. **IF** (*lower Lung* found) **OR** (*no Lung* AND *no Head*)  
    **THEN** Search *Pelvis + Hip* related features in bone related curves
  4. **IF** (*lower Lung* AND dataset continues caudally for at least 8 cm) **OR** (*Pelvis + Hip* found AND dataset continues cranially for min. 20 cm)  
    **THEN** Analyze *Liver* features (muscle dense and enhanced tissue)
  5. **IF** (*no Lung, Head or Pelvis + Hip* found AND scan range exceeds 40 cm) **OR** (*Pelvis + Hip* found AND dataset continues caudally for min. 20 cm) **OR** (*lower Lung* found AND at least 50 cm below diaphragm in dataset)  
    **THEN** Analyze *Knee* related features in dense tissue related curves
-

The respective steps per body region searched are explained in some detail below.

### *Lung*

The method starts with searching for the presence of normal lung tissue. An estimate of the lung volume is derived from the area per slice filled with lung dense voxels. In case this exceeds 250 ml and has an area with effective diameter of at least 8 cm on some slices we search for the presence of a steep increase/decrease of the lung fraction along the body axis starting from the slice depicting the maximum of lung tissue to determine positions near diaphragm or lung apex. If both are found we classify the presence of a complete lung and determine the lung center as mean of the two approximate end positions; otherwise, a partial lower respectively upper lung scan will be detected. If lung end positions can be determined the search for other body regions is greatly simplified. The presence of lung apex or diaphragm restricts the further search for other body region to the appropriate fraction of the scan. For example, there will be no pelvis above the diaphragm and no head below the apex. Furthermore, a scan depicting a certain length above/below the lung must contain certain other body regions in humans of common size even though specific features of those body parts cannot be identified in the computed families of curves.

### *Head and neck*

A head scan can be recognized from a ratio of spread in  $Y$  direction (patient anterior–posterior) to spread in  $x$  direction (patient right–left) exceeding 1.5 for the bony parts in the skull combined with a size estimate of the patient cross section below the area of a 20-cm disk. The base of the skull is commonly seen near the maximum of the gradient of a combination of the bony tissue fractions in the range above the lung. A position near the brain center is detected at the maximum of the soft tissue fraction above the base of skull. If any of the mentioned positions are found or the scan extends for some length above the lung apex the presence of neck or head is classified.

### *Pelvis and legs*

Views of the lower bones section show that pelvis and hip bone mainly consist of voxels with HU values between 100 and 500. Therefore, the classification analyzes the sum of the curve for “contrasted vessels” and “inner bone” that in most pelvis scans has close twin peaks, related to the pelvic bone and the hip position. To robustly find the twin peaks, the processed interval is limited to the section between diaphragm and lower end of the scan, respectively, a position somewhere in the thigh if present. This lower end of the search range in

the thigh is defined as highest point below the diaphragm where the effective diameter of the area filled with tissues above  $-300$  HU is below than 20 cm. A curve derived from effective diameter of all bone-dense areas is searched starting at the lower end of the search interval for the first local maximum exceeding the heuristically determined threshold of 10 cm. When one is found a search for another local maxima between 4 and 14 cm away from the first is performed. Depending on additional rules with heuristically determined parameters a differentiation of peaks related to hip and pelvic bone from peaks in the bone-related part of the statistical feature extraction present due to the use of stomach or colon CT contrast agent is attempted. The decision whether any and if which of the local maxima were related to hip and pelvic bone was implemented in two ways.

### *Liver*

A position near the center of the liver is identified from the maximum of the sum of areas filled by material in the categories non-enhanced and enhanced soft tissue and moderately contrasted vasculature. Such a maximum is only searched for below a position 4 cm above a mean diaphragm position and the lower end of the data, respectively, the pelvis position if this was identified in the step before.

### *Knees*

Presence of the legs at the lower end of the scan is assumed if the sum of dense material has an effective diameter below 20 cm. Within the leg range, starting at the upper end, a search for the highest local maximum of a smoothed curve of the sum of areas filled by material in the 100–500 HU range is performed. This search stopped at the knee joint in all datasets depicting the knees.

## **Evaluation**

### Performance and robustness

The method was implemented using the development platform MeVisLab [6] and tuned to successfully run on multiple collections of about 5,000 distinct CT scans. Therefore, key algorithmic components such as the background removal and the statistical feature extraction as well as the analysis of curve features were implemented as C++ MeVisLab module. The evaluation framework was implemented employing the MeVisLab graphical network programming approach and a scripting layer using python script and the MEVIS GUI definition language MDL. Evaluations were performed on state-of-the-art (as of 2007) Windows PC workstations, but also run on current Mac-OS and Linux systems due to

MeVisLabs cross-platform design. After algorithmic optimization run time was mainly determined by hard disk and network performance as the distributed datasets had to be accessed over the institute network.

The cases used during development evaluation mostly had more or less severe pathologies. Most came from patients with lung disease or final stage metastatic cancer and some, e.g., from patients with a scan of pelvis and the lower extremities for vascular disease.

However, since pathologies in most cases only affect a volume smaller than 100 ml they did not influence the statistical parameters substantially enough to cause problems. Even severe diffuse lung disease left most of the lung within the broad density range we considered as typical lung tissue.

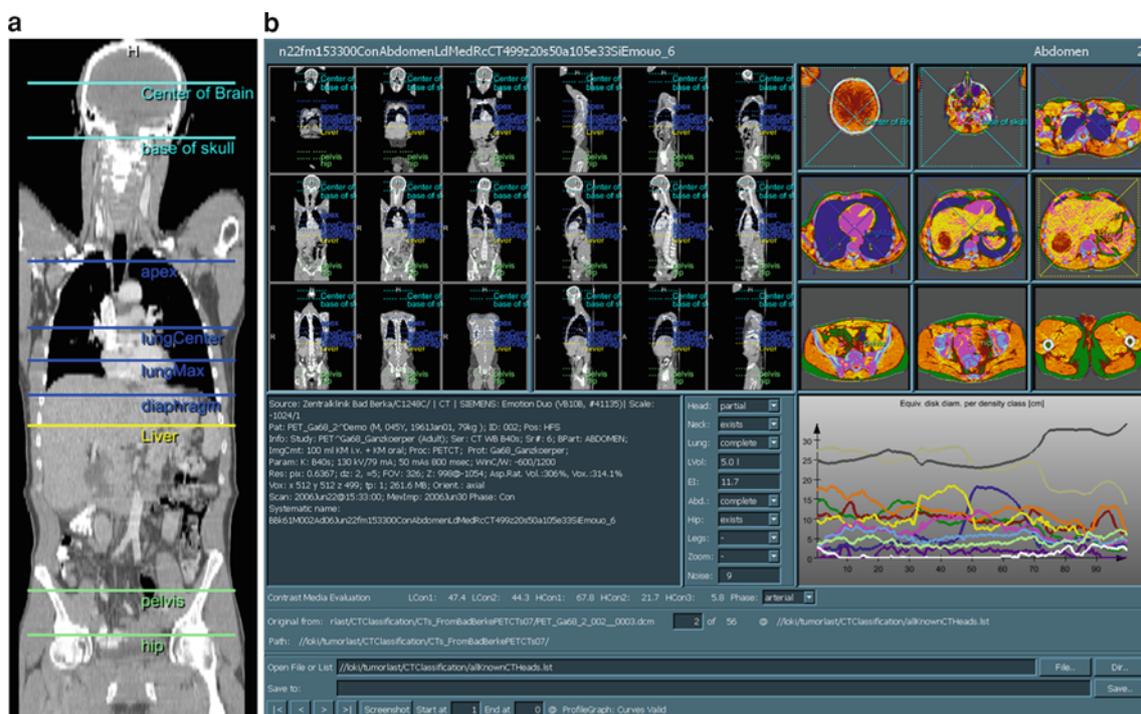
For each CT data set, the statistical density class features were derived with 3 mm resolution (or coarser with thicker CT slices) along the CT table and moderate local smoothing on datasets downsized with an in-plane down sampling factor of 2. In order to preserve the original image histogram shape as much as possible and to avoid long computations and loading, all slices over the network nearest neighbor resampling were employed along the Z axis. For the statistical feature extraction an in-plane sampling further reduced computations to every third voxel in every third line of an internally  $3 \times 3$  smoothed version these  $256 \times 256$  CT slices. This smoothing prior to voxel-wise classification was essential to become robust with respect to high noise. Effects related to

noise strongly depend on scan protocols (e.g. low dosage, high resolution) and patient size. Their influence on the classification was greatly reduced by the averaging employed prior to our voxel-wise classification of density values.

For each dataset a classification was derived, and a screenshot like the one in Fig. 4b was automatically generated. Including times for transferring the data over our network, resampling the data and generating multiple reformatted views in sagittal and coronal orientation from the resampled data for documenting the findings a processing rate of one screenshot per 11 s. was obtained on average. The density class feature extraction and classification itself accounted for only a small fraction (1–2 s) of the processing time.

## Verification

Correctness was confirmed by visual inspection of screenshots depicting the annotated sagittal, coronal, and axial views, and by judging the positions of the markers drawn for various features identified in the curve family analysis on the slice with the Z coordinate related to the positions where a feature was identified. Examples of such features are the position of max-lung-area and diaphragm: = highest position below max lung were lung area is below a percentage of the max lung area adjusted on a set of test cases (not part of the evaluation data base) low enough to yield a position usually near the middle of the range of slices depicting the



**Fig. 4** **a** (Left) Coronal view with indicators of characteristic positions determined from the curve families generated by the statistical feature extraction. **b** Example of a screenshot from a MeVisLab application for visual verification of the landmark position and classification results

caudal lung boundary but high enough to avoid classification of a position in the colon in, e.g., distended colonography scan as diaphragm. Other markers evaluated were, e.g., center of liver which was searched for over a combination of all material classes potentially found in abdominal scans of not extremely cirrhotic livers for CT studies performed native as well as with low or high doses of contrast agent on either arterial, venous, or late contrast phases.

Further possible marker positions are, e.g., center of brain, lung apex, the pair of peaks along the body axis in bony fractions related to pelvis and hip, and the knees. When during evaluation a marker was drawn on a slice depicting the respectively labeled region, the underlying feature recognition was considered correct.

We are convinced that the task of visual evaluation of the recognition of major organs and joints in CT scans does not require expert radiologist. Evaluation was therefore based on visual judgment by those authors sufficiently experienced in identification of the relevant structures in CT scans. Screenshots from nightly runs were stored under a systematic name derived from our classification results and basic DICOM information such as place and time of the image acquisition, patient birth year and sex, CT model name, geometry, and parameters of the reconstruction. With this approach collections of similarly classified data yielded screenshots that were grouped by classification result and allowed efficient evaluation of changes with respect to rated earlier runs. This allowed fast verification of the results (about 2–3 s per screenshot on average) because mismatches in the classification stood out from sets of similar screenshots. Thus, it became possible to visually check the classification findings of runs over up to 5,000 cases within a few hours per run, and to identify problematic cases needed to further refine the implemented rules and tune the internal parameters for a lower rate of false classifications. Tuning and refinement of the rules did not involve the data employed in the evaluation reported below.

#### Recognition rate and remaining problems

On the level of recognized depicted body regions a very high success rate was obtained. More than 98% of scans showing a complete body cross-section were correctly classified with regard to the imaged region. Some problems were identified in particular with cases where large foam positioning aids were not properly removed in the patient segmentation step. These were subsequently erroneously identified as equally dense lung tissue, e.g., at the height of head or feet in the image. Further limitations include zoomed details from the patient, e.g., reconstructions showing only a field of view (FOV) of about 15 cm around the heart, brain, spine, or major joints; however, due to their geometry they may still be correctly labeled as “zoomed”, and the common special

case of limited field of view heart studies can often be recognized from nearby lung on either side. On the level of marked positions along the body main axis results may still be improved in a substantial fraction of cases. Whereas lung-related markers are very reliable (only rarely erroneous due to corrupt data (e.g. missing slices) or failure in the patient contouring), positions derived from bony structures such as the markers for mandible, hip, or pelvis were sometimes a few centimeters beside their proper position. In the abdomen, this was often caused by the presence of colon contrast agent and most noticeable in large patients.

#### Evaluation of the noise level classification

The noise classification allows a very robust differentiation of images reconstructed for reading with focus on soft tissue findings (typical with soft tissue noise level as described above around 15 HU, at most 30 HU) and images reconstructed with high noise and spatial detail reconstructed with focus on findings related to lung and bones typical with a soft tissue noise level in the 40–80 HU range. The latter may also be differentiated from low-dose scans commonly used in lung cancer studies with a minimum soft tissue noise level exceeding 80 HU. So far, no wrong classifications with regard to the distinction between soft tissue and lung/bone reconstructions (including low dose scans) were observed. The noise levels were computed only over voxels of muscle density after moderate local  $3 \times 3$ -neighbor smoothing on images down-sampled by a factor of 2 using nearest neighbor resampling. The mentioned somewhat arbitrary thresholds were selected after initial experiments because they gave a very good separation of the described noise level classes that may also be identified in many cases from an inspection of the DICOM information through experts familiar with the various CT scanning parameters influencing the noise level (e.g. reconstruction kernel names, X-Ray dose and voltage as well as voxel size) and their coding in DICOM through different vendors. Our noise level classification indirectly also incorporates patient-specific aspects (in particular patient weight/size) influencing image quality that are not commonly found reliably in DICOM tags.

#### Unbiased evaluation on clinical trial data

The recognition rates were evaluated on a particular collection of 1,884 scans collected from an international multicenter oncological trial, provided by Bayer Vital, that was not used during development of the methods. The results were the following:

From the total collection 24 scans had missing slices filled with constant values or incorrect DICOM tags for scaling gray values to HU. Those had corrupted results from the statistical feature extraction and were discarded. Another 12

got no classification due to depicting only a very short scan range (5–7 cm) within the abdomen not providing sufficient context for the classification. This was in some cases caused by an erroneous splitting of a larger scan range into multiple short scans based on the available DICOM information during image import.

*Head* 43 (91.5%) of 47 scans showing the head were identified; the 4 remaining cases showed arms beside the head. This situation was not anticipated when implementing the rules and will require additional detection rules.

*Neck* All 28 CT scans (100%) covering the neck region were correctly identified.

*Lung* 1,771 (99.7%) of 1,775 scans containing a substantial part of the lung were correctly classified as either upper, lower, or complete lung. The four failures were due to pre-processing problems with the slice-wise background removal obscuring the statistical features at certain positions.

*Liver* On 1,536 (99.5%) of 1,543 scans depicting most of the liver, a position near the liver center was correctly marked.

*Abdomen* The presence of the abdominal region, defined as the body part between the center of the liver and a position between the lower end of the kidneys and the highest point of the pelvic bone, was correctly identified in 936 (97.7%) of 958 scans ranging over some part of the abdomen. An attempted distinction between scans rated as upper abdomen only and full or lower abdomen was not very robust; however, such a distinction also lacks a clear anatomical definition.

*Pelvis* Detecting the pelvis region in the density class feature curves was well possible but the least reliable part in our experiments. The search for twin peaks in the bony area curves correctly identified 955 (95.8%) of 997 scans showing the pelvis and hips with the so far most successful implementation of an admissible peak pair detection algorithm. The algorithm searches two peaks with plausible values for (a) bony and soft tissue area for a hip or pelvis cross-section (b) distance to lung if present, and (c) distance between pelvis and hip peaks. However, the positions marked as center of the pelvis and hip joint were somewhat variable over similar scans or between variants of the algorithm and frequently not at the proper location. Most of the failures were due to colon contrast agent mistakenly considered as pelvic bone, especially in large patients.

*Lower extremities* Legs and knees were not scanned in the oncological trial. During development, knees were detected in all arterial run off and PET/CT data presented to the algorithm whenever the knees were in the scanned range.

## Discussion

### Problems

A crucial step in our method is the patient contour determination. Its failure seldom leads to inclusion of foam material near the head or legs, or partial removal of the lung in zoomed thorax data resulting in absurd lung end positions and consequently useless results.

A principle problem are CT data that have erroneous DICOM information about rescaling gray level information to Hounsfield units or data with some missing slices filled by constant values yielding useless results of the statistical feature extraction.

Furthermore, scans over only a short range frequently fail to obtain a proper classification due to missing context. Attempts to classify the individual slices based on features of the statistical feature extraction may help to improve the results, but initial work on classification without the use of context information found no sufficiently robust features so far to recognize the depicted region from single slices based on the computed statistical parameters alone. More features or an analysis using advanced classification based on rules extracted from training sets rather than implemented rules derived from human anatomical knowledge may improve this approach. However, we believe this would require very large training sets to cover the board variability seen in body shapes and different uses of CT contrast material.

The principal challenge to our method is provided by scans depicting only a ROI inside the patient (e.g. near heart, spine or major joints, one extremity only). These may be recognized as zoomed scans simply from the small depicted area not showing a head, but detailed classification in the spirit of our approach remains a challenge as many different cases are possible. A reasonable recognition rate was so far only achieved with special treatment of zoomed data for certain common types of cardiac scans.

Another current limitation is the lacking of support for non axial data becoming clinically more frequent with the spread of Multislice-CT scanners. This limitation may be overcome at some computational cost by reformatting the data prior to analysis or a straightforward alternative implementation of the statistical feature extraction in case of coronal- or sagittal-reconstructed CT data. Such scans are easily recognizable as reformatted data from reliable geometric DICOM information on image orientation. Oblique reformatted CT data are possible. These would require a more refined approach, but will remain rare in clinical use.

The exact location of a determined position outside the lung will vary sometimes by a few centimeters between different reconstructions of the same CT scan, between repeated scans of the patient, or when considering a sub-range of the scan rather than the full scan. Further results might differ

after minor modification of algorithmic parameters. Sometimes certain positions are correctly detected in one scan but not in a very similar one of the same patient. Nevertheless, the obtained overall classification of the depicted body regions remains the same due to some redundancy in the classification in the great majority of cases. More robustness of the detection of non-lung landmark positions to consistently separate body regions along the *Z*-axis may require more elaborate image analysis for more specific landmarks. Such analysis may be performed efficiently when restricted to a small ROI that can be obtained from the regions identifiable from the statistical features discussed above.

## Conclusion

We believe the presented method provides a robust and novel concept for a fast, yet reliable coarse level classification of CT databases. This holds the potential to foster evaluation of various algorithms based on substantial case numbers needed to obtain a confidence in the robustness of image processing methods required for their implementation in clinical use. Furthermore, the classification may aid post processing or reading software to perform its task with algorithmic or visualization parameters optimized with respect to the present patient's size, image noise level, and contrast use. In addition, it might help to establish standardized basic

quantification, like an air-lung ratio most likely correlating with severe emphysema or an abdominal body fat ratio that may be automatically computed at almost no cost on all scans where the computation is appropriate from the density class feature curves. Such quantifications may be employed, e.g., to retrieve cases for an evaluation of more refined emphysema quantification methods.

## References

1. Greenspan H, Goldberger J, Ridel L (2001) A Continuous probabilistic framework for image matching. *J Comput Vis Image Underst* 84(3):384–406
2. Pinhas A, Greenspan H (2004) A continuous and probabilistic framework for medical image representation and categorization. In: *Proceedings of SPIE*, vol 5371, pp 230–238
3. Lehmann T, Wein B, Keysers D, Kohnen M, Schubert H (2002) A mono-hierarchical multi-axial classification code for medical images in content-based retrieval. In: *Proceedings of the 1st IEEE International Symposium on biomedical imaging*, pp 313–316
4. Deselaers T, Müller H, Clough P, Ney H, Lehmann TM (2007) The CLEF 2005 automatic medical image annotation task. *Int J Comput Vis* 74(1):51–58
5. Zhou X, Kamiya N, Hara T, Fujita H, Yokoyama R, Kiryu T, Hoshi H (2004) Automated Recognition of Human Structure From Torso CT Images, VIIP. In: *International Conference on Visualization, imaging and image processing*. Marbella, Spain
6. MeVisLab, Software for Medical Image Processing and Visualization. <http://www.mevislab.de>