Integral Geometry Descriptors for Characterizing Emphysema and Lung Fibrosis in HRCT Images

Michal Charemza[1], Elke Thönnes[1,2], Abhir Bhalerao[3], and David Parr[4]

[1] Centre for Scientific Computing, University of Warwick, UK,
m.t.charemza@warwick.ac.uk,
[2] Department of Statistics, University of Warwick, UK, e.thonnes@warwick.ac.uk,
[3] Department of Computer Science, University of Warwick, UK,
abhir.bhalerao@dcs.warwick.ac.uk,
[4] Department of Radiology, University Hospital Coventry and Warwickshire NHS
Trust, Coventry, UK, david.parr@uhcw.nhs.uk

Abstract. Integral geometry descriptors are used to characterize local textural properties of lung parenchyma from HRCT images. These quantities, known as Minkowski functionals, describe the morphology and topology (connectivity) of 2D and 3D binary structures. They have been shown to be effective in describing properties of complex and disordered media, such as open foams. In this paper we describe the Minkowski functionals and propose their use for detecting and grading emphysema and fibrosis in HRCT images. We show which relevant combinations of the four quantities correlate well with three degrees of severity of emphysema and states of fibrosis. We present some illustrative results and make proposals for the use of these descriptors in a larger validation study.

1 Introduction

Emphysema and pulmonary fibrosis are common respiratory disorders which destroy the lung and reduce its ability to oxygenate blood. The primary cause is smoking, but other more subtle risk factors, such as genetic predisposition, are known to be involved. In terms of health care costs, it ranks among the top five western world diseases. The development of robust quantitative methods for its early diagnosis, monitoring of its treatment and understanding of its disease process is therefore of great importance.

High Resolution Computed Tomography (HRCT) is used regularly to assess lung function and structure, and has become an essential tool in the detection and assessment of emphysema and related conditions such as lung fibrosis. The disease is characterized by the destruction of the lung's alveolar sacs and the collapse of their walls which results in localized fibrosis. This leads to a corresponding loss of the respiratory capacity of the lung.

In HRCT images, the loss of lung tissue manifests itself as a reduction of the mean lung density, which can be measured. One simple way is to locate

lung regions below a given low threshold value (e.g. -910 Hounsfield Units) for patients holding their breath at full lung capacity [1]. Histogram based analysis has also been proposed to better determine a cut-off threshold. Nevertheless, the use of a single, global density index has been shown to be unreliable during early or mild stages of the disease, or when the emphysematic regions are dispersed in normal parenchyma [2]. Also, when the pathologic process is mixed, such as in the presence of inflammation and fibrosis, the local density can actually increase. Emphysema and fibrosis can create complex patterns of the lung, whose radiological appearance is commonly known as ground-glass opacities (GGO), honeycombing (HC) and irregular linear or reticular infiltrates [3].

Since the appearance of the disease progression is *textural*, it is unsurprising that image texture analysis techniques have been applied to the problem of robust quantification [2, 4, 1, 5, 3]. The early work of Uppaluril et al. [2] used first and second-order texture features together with fractal dimension (see [6]) to characterize emphysema from CT images. Using pattern recognition, they performed feature selection and trained a classifier to achieve an accuracy of approximately 90% and showed it to be significantly better than mean lung density (MLD) and histogram analysis. Chabat et al. [4] proposed the use of grey-level co-occurrence matrix (GLCM) texture measures to produce a 13 dimensional feature vector containing common features of the GLCM, such as energy and entropy. They also included a number of local shape features by considering the number of primitives at a given grey level (quantized to 16), of a predefined radius. These were then used to estimate the "emphasis", size (short and long) and uniformity, which would capture the connectivity and size of the primitives. This approach is empirical but could be interpreted in the integral geometry framework being presented here. The use of GLCM and the related shape and connectivity measures suggested by Chabat are used again in more recent work of Xu et al. [5]. This time, 3D voxel neighbourhoods were considered and the classifier was set up to grade the pathology into severe, mild and normal lung appearance. Across 34 subjects, they achieved significantly better discrimination over an equivalent 2D approach. In reference [1] and recently Zavaletta et al. [3], spatial maps of lungs have been produced which graphically show the results of the texture classification and can be visually related to the degenerative states of the disease. The type of classifiers used vary from simpler kNN ones, as used in [3], to Bayesian [5] and to neural network approaches [7].

Other work which is of relevance is the study of distributions of air-spaces in the parenchyma and their progression [8], something which could be readily estimated by integral geometry descriptors. Notably, in-vitro histopathological analysis of lung architecture is reported by Ochs et al. [9]. They calculated Euler-Poincaré characteristics from histological samples and were able calculate the total number of alveoli in the lung based on an estimate of the shape coefficient of the alveolus. In this paper, we propose the use of integral geometry descriptors: the Minkowski functionals, with the aim of producing calibrated maps of the emphysema and related fibrosis, localizing and grading the lung parenchyma into several stages from normal to severe. After a brief introduction to the math-

ematical foundations behind Minkowski functionals, we present experiments on synthetically generated data (based on a binary point-process model), and axial HRCT slices of a subject. We discuss the results finding and make proposals for further work.

2   Integral Geometry Descriptors

Healthy and emphysematic lung tissue present textures of distinct morphological properties. To describe these we need geometrical and topological characteristics that describe not only volume and shape but also connectivity. Integral geometry provides a suitable family of such descriptors: *Minkowski functionals or intrinsic volumes*. In 3D there are four such functionals and these are proportional to more commonly known quantities such as volume, surface area, mean breadth and the Euler-Poincaré characteristic.

Intrinsic volumes have been successfully used in material science to characterize and discriminate morphology of various media [10, 11]. In material science the structure of interest is usually homogeneous. In this paper we are interested in characterizing a medium that is composed of various locally homogeneous but texturally differing regions. To do so we compute Minkowski functionals for equal-sized regions of voxels rather than the whole image.

Minkowski functionals are unbiased, stereological estimators and provide local and global morphological information. In contrast to more standard characterisation methods such as pair-correlation functions or cord-length distributions, they incorporate information from higher order correlations. Minkowski functionals have nice mathematical properties such as C-additivity, see (4), which means that they can be computed efficiently based on simple configuration counts.

Mathematically, the Minkowski functionals are defined for a convex, compact set $K \subset \mathbb{R}^3$ via Steiner's formula. Let $K \oplus B_r$ be the dilation of the set K by a closed ball of radius r centered on the origin. Then the volume V of $K \oplus B_r$ can be written as a polynomial function of r as follows:

$$V(K \oplus B_r) \quad = \quad \sum_{k=0}^{3} \binom{3}{k} W_k(K) r^k. \tag{1}$$

Here W_k is the kth Minkowski functional. For example, if C is a cube of side-length a then

$$V(C \oplus B_r) \quad = \quad a^3 + 6a^2 r + 3a\pi r^2 + \frac{4\pi}{3} r^3 \tag{2}$$

and so $W_0(C) = a^3, W_1(C) = 2a^2, W_2(C) = a\pi$ and $W_3(C) = 4\pi/3$. This also illustrates the relation of between the Minkowski functionals and the more common descriptors volume V, surface area S, mean breadth B and Euler-

Poincaré characteristic χ:

$$W_0(K) = V(K), \qquad W_1(K) = \frac{1}{3}S(K), \tag{3}$$

$$W_2(K) = \frac{2}{3}\pi B(K), \qquad W_3(K) = \frac{4}{3}\pi \chi(K).$$

As the reader is more likely to be familiar with the latter quantities, we will present our results in terms of V, S, B and χ. The Minkowski functionals are proportional to these quantities and so we simply refer to these descriptors as Minkowski functionals.

3 Method

To compute a Minkowski functional for a binary image their definition is extended from compact convex sets to sets that are finite unions of such. This is done exploiting C-additivity of Minkowski functionals, that is, for compact convex K_1 and K_2 we have

$$W_k(K_1 \cup K_2) \quad = \quad W_k(K_1) + W_k(K_2) - W_k(K_1 \cap K_2). \tag{4}$$

Further their definition is extended to the interior A_n° of a compact n-dimensional geometrical object A_n in 3D space as follows [12]:

$$W_k(A_n^\circ) \quad = \quad (-1)^{3+n+k}W_k(A). \tag{5}$$

These extensions are required as each voxel is assumed to be a cube of edge length 1, and the image is defined to be a number of white voxels in regular lattice of a black background. Each white voxel is subdivided into its component geometrical objects, the interiors of which are disjoint. Using (4) and (5), the computation of the Minkowski functionals now reduces to the problem of counting these objects: i.e. finding the numbers of open cubes n_3, open faces n_2, open edges n_1, and vertices n_0 ensuring that for adjacent white voxels each shared geometrical object is only counted once.

By repeated application of (4) and (5), and using the known values of the Minkowski functionals for a cube, face, edge and vertex, the Minkowski functionals can then be shown to be [12]:

$$V = n_3,$$
$$S = -6n_3 - 2n_2,$$
$$B = \frac{3}{2}n_3 - n_2 + \frac{3}{2}n_1,$$
$$\chi = -n_3 + n_2 - n_1 + n_0. \tag{6}$$

For our data, to find n_3, n_2, n_1 and n_0, we use the method in [13], referred to as the *algorithm of equations*. A black voxel does not give any contribution to the sums of geometrical objects. Given a white voxel we find its contribution to

the total of geometrical objects, not counting any shared geometrical objects of thirteen of the twenty-six adjacent voxels. Indeed, Blasquez and Poiraudean [13] refer to these voxels as the *preceding* voxels, however their contributions are independent of the order in which they are found.

4   Experiments

To analyse a binary image, the image was split into a number of identical cuboids of a given size in order to describe the heterogeneity of the data. Each cuboid was then embedded in a black background, to ensure the boundary geometrical objects are all counted in a consistent manner. Then the Minkowski functionals were found for each cuboid and maps of the Minkowski functionals were produced. We investigated the correlation of the 4 quantities by principal component analysis (PCA) and produced grey-scale maps of the principal variation and the first three principal components mapped to the green, red and blue channels respectively.

4.1   Synthetic Data

We first use the described method to test whether Minkowski functionals can differentiate between idealized, and fairly crude, models of diseased and healthy lung tissue.

The models are sampled in a window of size $320 \times 500 \times 30$ containing 3 partially overlapping regions of equal overlaps. The first two models are Boolean models [14] which use a random point process to place different structuring primitives or grains, see Figure 1. The third model is a Voronoi tesselation [14]. In a Boolean model the locations of grains are independent and identically distributed uniformly on the region. For the healty tissue, the grains are $2 \times 2 \times 2$ cubes and 3500 grains are placed in total. For the fibrotic tissue, 4000 points are chosen, and at each point is placed a cuboidal segment. The segments have lengths that are uniformly distributed on the interval $[15, 20]$, and gradients that vary between ± 0.8. Their thickness is fixed and is 5 in the z-axis and 2 in the xy-plane. Finally, for the emphysematic (honeycombed) tissue 800 points are placed at random and the corresponding Voronoi tesselation is computed. The thickness of the walls of each Voronoi cell is chosen to be one voxel. Figure 1 shows slice 15 of the synthesized volume.

We produced a scatter plot of the 4 functionals projected onto the first three principal components of the covariance of the output, see Figure 2. Data from the five vertical bands of the synthetic image are plotted and clearly show the distribution in the Minkowski sub-space and good separation of the different structures. We used the same dimensions to map the descriptors onto the green, blue and red colour channels to produce the descriptor map in Figure 1 (bottom-right).
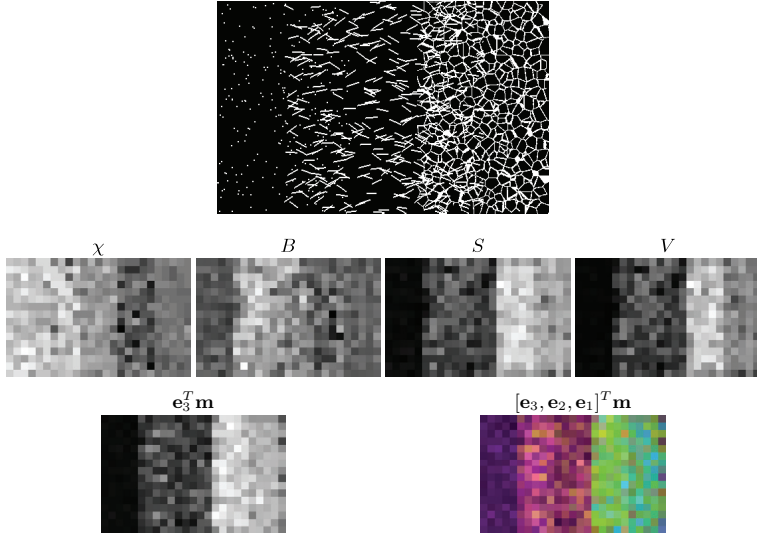
χ B S V

$\mathbf{e}_3^T\mathbf{m}$ $[\mathbf{e}_3,\mathbf{e}_2,\mathbf{e}_1]^T\mathbf{m}$

Fig. 1. Minkowski functionals found on synthesized data generated using Boolean and Voronoi models. Top: Slice 15 of synthesized volume size $320 \times 500 \times 30$ (parameters of models are given in section 4.1). Second row: Minkowski functionals found on voxel regions of size $20 \times 20 \times 10$ from slices 11 to 20. Bottom row: Projection of functionals onto principal mode of variation $\mathbf{e}_3$ from PCA of data $\mathbf{m} = (V, S, B, \chi)^T$; RGB mapping of three principal modes of variation.

4.2   HRCT Lung Images

An axial HRCT scan of a subject with lung fibrosis was acquired using a GE Lightspeed Plus CT scanner at a tube current of 150mA. The slice images were of size 512×512 with in in plane resolution of 0.684×0.684 mm and a slice thickness of 0.625 (Figure 3). We calculated the Minkowski functionals on thresholded lung data, using a threshold of -600 Hounsfield units, and taking voxel regions of size $6 \times 6 \times 4$.

In order to visualize the results, we performed principal component analysis on the output from axial slices 29 to 33, and produced colour descriptor maps by mapping the 4 measures on to a 3D sub-space spanned by the first three principal axes of the PCA. Illustrative results for three slices are shown in Figure 4. The colours are, approximately: normal tissue appears as green; fibrosis as a blue to purple hue and honey-combing as brown and red.
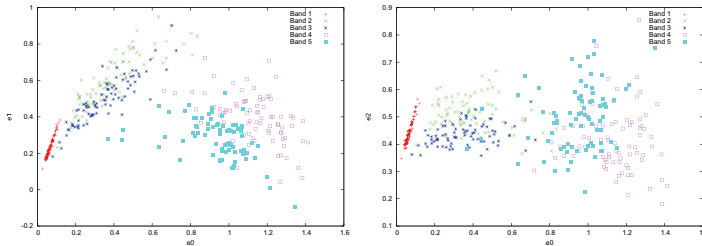
Fig. 2. Scatter plot of measures from vertical bands of synthetic image projected onto pairs of the first thee principal components of variation, $[\mathbf{e}_0, \mathbf{e}_1]$ and $[\mathbf{e}_0, \mathbf{e}_2]$. Bands are numbered from left-to right (with respect to images in Figure 1).
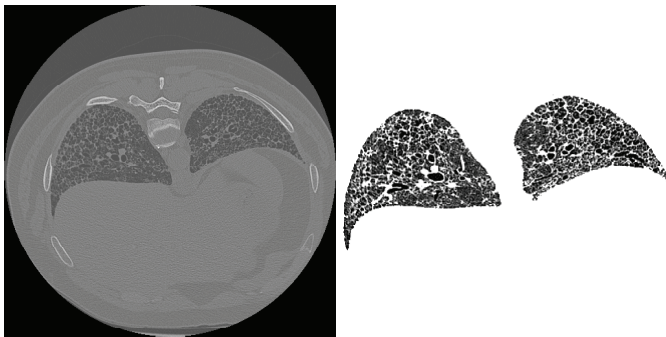

Fig. 3. Original and close-up of 'stripped' data of axial slice 40, size 512×512, with in-plane voxel resolution of 0.684×0.684mm. This slice clearly shows the honeycombing artefacts and fibrosis.

5   Discussion and Conclusions

The results of Figure 4 show that Minkowski functionals can discriminate between different stages of degeneration in emphysematic/fibrotic lung. Minkowski functionals not only have the advantage of being descriptors that have a clear geometric/topological interpretation, see Table 1, but are also easy and fast to compute. They are fully 3D descriptors that go beyond simple local lung density estimation.

After submitting this paper we discovered concurrent but independent work on the use of Minkowski functionals for grading of emphysematic/fibrotic lung by Boehm et al. [15, 16]. The authors select specific volumes of interest that have been labelled by experts as either normal, emphysematic or fibrotic lung tissue. The Minkowski functionals of these volumes are then computed as a function of

Table 1. Interpretation of Minkowski functionals for homogeneous structures. If W is the volume of the sampling window then $V_v = V/W, S_v = S/W, B_v = B/W$.

Specific fibre length (edge density) L_V	$B_v/\pi(1 - V_v)$
Number of particles	Isolated particles - χ
Mean thickness of edges	$S_v/\pi L_v$
Mean section area of edges	V_v/L_v
Porosity (volume fraction) p	$1 - V_v$

the threshold value used in the binarisation of the image. Using an integrative filtering procedure a single numerical quantity is produced that then can be used for classification. As in our study, the authors show that Minkowski functionals are effective tools for the grading of emphysematic/fibrotic lung and provide information that differs from densitometric measures. In contrast to Boehm et al. we do not classify volumes of interest, but rather aim at the segmentation of the whole lung into the various stages of degeneration. This is achieved by using localized versions of the Minkowski functionals that can then be displayed as spatial maps, see Figure 4. In further work we are planning to examine distribution of these localized functionals for the various stages of pathology. Modelling the spatial distribution of the localized topological measures as a mixture distribution then provides the basis of segmentation of the lung.

In Section 4.1 we present a crude model for some of the structures that may be encountered in diseased lung. More advanced probabilistic model development is needed as important groundwork for statistical inference and the analysis of longitudinal data. The computation of Minkowski functionals is a first step towards model fitting as explicit expressions of these functionals are known for various models, for example the Boolean model.

While the promise of the proposed descriptors is clearly illustrated with our results, supervised validation on a larger survey is needed and planned as further work. In this paper, a simple PCA illustrates how the Minkowski functionals differentiate between textures. Further work will exploit these characteristics in more sophisticated method for feature selection [1] as well as classification and learning [7].

Acknowledgments

References

1. Hoffman, E.A., Reinhardt, J.M., Sonka, M., Simon, B.A., Guo, J., Saba, O., Chon, D., Samrah, S., Shikata, J., Tschirren, J., Palagyi, K., Beck, K., McLennan, G.: Characterization of the Interstitial Lung Diseases via Density-Based and Texture-Based Analysis of Computed Tomography Images of Lung Structure and Function. Academic Radiology **10** (2003) 1104–1118
2. Uppaluri, R., Theophano, M., Sonka, M., Hoffman, E.A., McLennan, G.: Quantification of Pulmonary Emphysema from Lung Computed Tomography Images. American Journal of Respiratory and Critical Care Medicine **156** (1997) 248–254
3. Zavaletta, V.A., Bartholmai, B.J., Robb, R.A.: High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis. Academic Radiology **7** (2007) 769–771
4. Chabat, F., Yang, G., Hansell, D.: Obstructive lung diseases: Texture classification for differentiation at CT. Radiology **228** (2003) 871–877
5. Xu, Y., Sonka, M., McLennan, G., Hoffman, E.A.: MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. IEEE Transaction on Medical Imaging **25** (2006) 464–475
6. Shimizu, K., Johkoh, T., Ikezoe, J., Ichikado, K., Arisawa, J., Nakamura, H., Tamura, S., Nagareda, T.: Fractal analysis for classification of ground-glass opacity on high-resolution CT: an in vitro study. Journal of Computer Assisted Tomography **21**(6) (1997) 955–961
7. Friman, O., Borga, M., Lundberg, M., Tylén, U., Knutsson, H.: Recognizing Emphysema - A Neural Network Approach. In: Proceedings of 16th Int. Conf. on Pattern Recognition - ICPR'02. (2002) 10512
8. Nakano, Y., Saiki, H., Muro, S., Hirai, T., Oku, Y., Nishimura, K., Mishima, M.: Comparison of low attenuation areas on computed tomographic scans between inner and outer segments of the lung in patients with chronic obstructive pulmonary disease: incidence and contribution to lung function. Thorax **54** (1999) 384–389
9. Ochs, M., Nyengaard, J.R., Jung, A., Knudsen, L., Voigt, M., Wahlers, T., Richter, J., Gundersen, H.J.G.: The Number of Alveoli in the Human Lung. American Journal of Respiratory Care Medicine **169** (2004) 102–124
10. Ohser, J., Mücklich, F.: Statistical Analysis of Microstructures in Materials Science. Statistics in Practice. John Wiley & Sons, New York (2000)
11. Arns, C., Knackstedt, M., Mecke, K. Lecture Notes in Physics 600. In: Charactersing the Morphology of Disordered Materials. Springer (2002)
12. Michielsen, K., De Raedt, H.: Integral-geometry morphological image analysis. Physics Reports **347** (July 2001) 461–538
13. Blasquez, I., Poiraudeau, J.F.: Efficient processing of Minkowski functionals on a 3D binary image using binary decision diagrams. Journal of MSCG **11**(1) (2003)
14. Stoyan, D., Kendall, W., Mecke, J.: Stochastic Geometry and its Application. Probability and Statistics. John Wiley & Sons, New York (1995)
15. Boem, H., Fink, C., Becker, C., Reiser, M.: Automated Characterization of Normal and Pathologic Lung Tissue by Topological Texture Analysis of Multi-Detector CT. In Giger, Karssemeijer, eds.: Medical Imaging 2007: Computer-Aided Diagnosis, Proceedings of SPIE Vol 6514. (2007) DOI: 10.117/12.702697
16. Boehm, H., Fink, C., Attenberger, U., Becker, C., Behr, J., Reiser, M.: Automated classification of normal and pathologic pulmonary tissue by topological texture features extracted from multi-detector CT in 3D. European Radiology **18** (2008) DOI: 10.1007/s00330-008-1082-y.
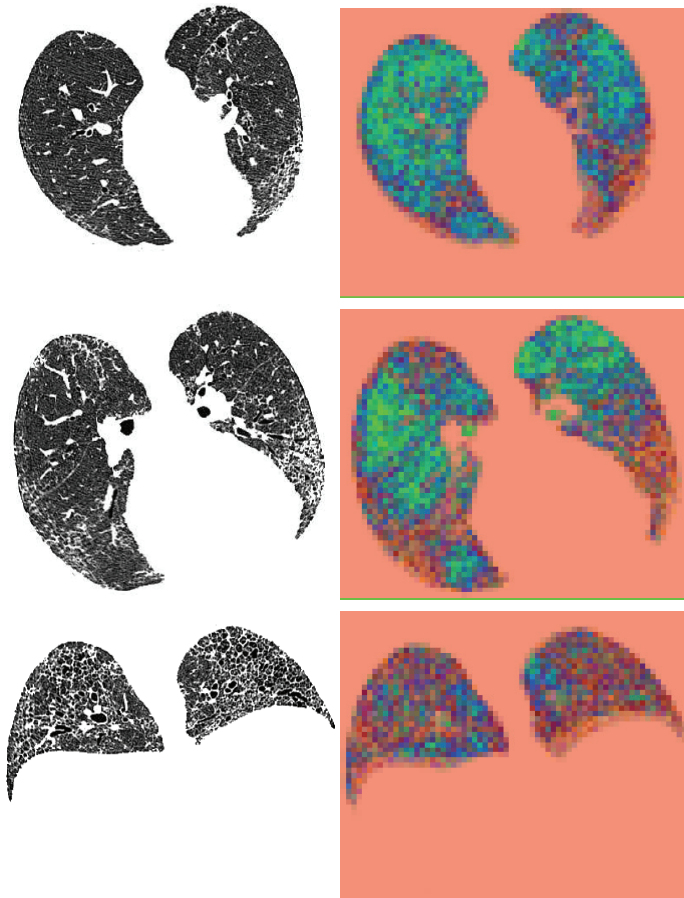
Fig. 4. Results on axial HRCT slices of subject with severe pathology. Left column: original HRCT data. Right column: Results of PCA projection of first 3 principal modes of variation from covariance analysis of slice 30 (second row). Slices 20, 30 and 40 shown. Normal tissue colours as green; fibrosis colours as a blue-purple hue; honey-combing as brown and red.