UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
Electronics and Computer Science

Volumetric texture analysis based on three dimensional Gaussian Markov random fields

by

Yasseen Hamad Al Makady

Thesis for the degree of Doctor of Philosophy

17th September 2020

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES
Electronics and Computer Science

Doctor of Philosophy

VOLUMETRIC TEXTURE ANALYSIS BASED ON THREE DIMENSIONAL GAUSSIAN MARKOV RANDOM FIELDS

by Yasseen Hamad Al Makady

The analysis of rich information provided by volumetric data is paramount for developing robust applications, especially in the medical domain where volumetric images are common. The volumetric texture is a fundamental characteristic that plays a vital role in understanding and analysing volumetric images. The current approaches to volumetric texture analysis commonly extract texture features from two-dimensional (2D) slices of the volumetric images while ignoring the valuable information provided by the third dimension in volumetric data. This approach is therefore not capable of efficiently analysing textures in emerging three-dimensional (3D) images, which results in information loss. Consequently, the development of new methods to analyse volumetric texture is essential because of recent advances in 3D technologies of imaging, especially medical imaging.

Among texture analysis methods, model-based Gaussian Markov random fields (GMRFs) are emerging as a choice for modelling texture and have been used to characterise textures in 2D images. Extending the use of GMRF to characterise textures in volumetric images has not received much attention, and therefore exploring this would be beneficial in many applications, especially in medical image analysis applications in which rich texture information is available in the form of volumetric images.

This thesis proposes a new method based on GMRF to characterise textures in volumetric images. The 3D GMRF is developed to extract texture features from 3D patches of volumetric images taking into account the information found in the third dimension. The features extracted by the proposed method are then employed for volumetric texture classification, segmentation and texture-based region tracking. The challenges that arise while dealing with textures in 3D images, such as the growth of sampling points, achieving rotation invariance, and the high dimensionality of feature vectors, are investigated and appropriate solutions are proposed.

The proposed methods demonstrate a higher performance compared to other methods through a comparison evaluation carried out on a synthetic textures database. Moreover, the methods proposed here are exploited to solve three real-world problems by attempting to diagnose emerging serious chronic obstructive pulmonary disease (COPD), classifying subjects with lung cancer, and analysing cilia motion using clinical datasets. The proposed methods do not require extensive training data or powerful hardware, which makes them suitable for medical image applications where datasets are usually small.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Yasseen Hamad Al Makady, declare that the thesis entitled *Volumetric texture analysis based on three dimensional Gaussian Markov random fields* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published, listed as [Publications P1-P6] under Section 1.5.

Signed:..

Date:...

Acknowledgements

I would like to express my sincere gratitude and appreciation to everyone who supported and encouraged me throughout my PhD journey. First and foremost, I would like to thank my supervisors Dr. Sasan Mahmoodi and Prof. Mark Nixon for their dedicated support, supervision, and guidance throughout this research. I am also thankful to Dr. Michael Bennett for providing me with his views and constructive feedback. I would like also to thank my examiners, Professor Sir Michael Brady and Dr Christine Evers, for their constructive feedback and valuable comments. Further, I would like to thank the Technical and Vocational Training Corporation (TVTC) in Saudi Arabia and the Saudi Arabian Cultural Bureau in London for financially supporting my PhD scholarship. Finally, I extend my gratitude and thanks to my parents, brothers, and sisters for their continued love and encouragement. Many thanks also to my wife, who has been constantly motivating me and keeping me optimistic during this period.

To my sons, Ghassan and Abdulaziz.

Chapter 1

Introduction

1.1 Context

The presence of texture in our natural world demonstrates its importance in the characterisation of objects, exhibiting rich information. In computer vision, texture provides important visual cues about objects' appearance and properties. It plays a key role in many applications in computer vision, including remote sensing, medical image analysis, object recognition, and automated inspection.

Volumetric texture (or 3D texture) is the texture that can be found in 3D images and is indexed by $(x, y, z) \in \mathbb{R}^3$. This type of texture has become increasingly available in the medical imaging domain due to the advancement of data acquisition techniques in medical imaging, such as magnetic resonance imaging (MRI) and high-resolution computed tomography (HRCT). Such techniques allow the acquisition of textured volumetric images that leverage the rich information of the organ's internal structure. Analysing volumetric texture is commonly carried out by extracting textures from 2D slices of volumetric images. However, this approach prevents sufficient exploitation of the valuable information provided by the third dimension in volumetric data.

In medical applications, texture analysis gained its own importance in the diagnosis of diseases based on a large amount of textures contained within medical images. Many diseases have symptoms that appear as textural changes in the medical images of soft tissue. However, these changes in the texture of images are difficult to analyse using human perceptions, especially for some diseases such as cancer, in which normal cells are very similar to abnormal cells, leading to inaccurate diagnoses (Faust et al. 2018). Therefore, volumetric texture-based methods are essential for medical image analysis to provide reliable, fast and accurate diagnoses.

This thesis aims to develop a robust method that exploits the rich information found in volumetric texture and employ it to detect different diseases. This involves extracting rotation invariant features and using them for texture classification. We also propose a method for volumetric texture segmentation to detect the tissues affected with a specific disease. Furthermore, the features extracted by our method are also exploited to track the textured region that characterises some diseases.

1.2 Contribution

The primary contribution of this research is to develop a three-dimensional Gaussian Markov random fields (3D GMRF)-based method to extract volumetric texture features from 3D images rather than analysing texture by using the 2D approaches. Consequently, we highlight the following contributions:

1. **Extension of GMRFs to 3D GMRFs.** The extension of GMRFs to 3D GMRFs for the purpose of extracting texture features from volumetric images is introduced. The 3D GMRF model parameters are locally estimated to extract volumetric texture for each voxel. This estimation is performed by using an estimation cub and 3D neighbourhood system. The use of spherical neighbourhood with an equidistant spherical sampling scheme enables exploration of the rich information found in volumetric texture.

2. **3D GMRF-based binary features for volumetric texture classification.** Various novel descriptors based on features extracted by the 3D GMRF model are introduced for volumetric texture classification. This includes improved descriptors by encoding the extracted features into single codes using a binary encoding scheme before their distribution over the entire volume is computed. It is demonstrated that employing the binary encoding strategy for 3D GMRF parameters improves classification performance with less computation time.

3. **Novel rotation invariant descriptor**. We introduce a novel rotation invariant descriptor based on spherical harmonics for volumetric texture classification. The descriptor demonstrates improved classification performance compared to other state-of-the-art 3D texture descriptors. The two main advantages of this descriptor are that it is not influenced by increasing the number of sampling points, and it does not require a large dataset for training.

4. **New adaptive method for volumetric texture segmentation.** The texture features extracted by 3D GMRF are used for volumetric texture segmentation. We develop an adaptive method that overcomes the parameter selection during estimation and filtering stages. This method selects the optimal settings for the estimation cube and averaging filter to improve the segmentation performance.

5. **Texture-based region tracking**. We introduce a novel texture-based tracking method to track a textured region and extract its trajectory. Instead of tracking

individual pixels, groups of pixels that represent the moving texture are tracked based on the features extracted by 3D GMRF.

1.3 Applications

The method developed in this research is employed as a computer-aided utility for the following problems:

1. **Diagnosis of chronic obstructive pulmonary disease (COPD).** The rotation invariant descriptor is employed for COPD detection and classification, which performed in a clinical dataset of 32 subjects, of which 19 are healthy and 13 have been diagnosed with COPD.
2. **Lung nodule classification**. The rotation invariant descriptor is also employed for benign and malignant lung nodule classification using the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).
3. **Lung segmentation.** The developed method is utilised to localise the regions of the lung that are affected by COPD.
4. **Ciliary motion analysis.** The proposed method is successfully utilised to extract the trajectory of the ciliary motion, which could help to analyse the beating behaviour of cilia.

1.4 Datasets used in this research

The methods developed in this research are evaluated on various datasets. These are as follows.

1. **The RFAI database.** This is a synthetic database for 3D texture images proposed by Paulhac, Makris, Ramel et al. (2009). This database is used in this research to evaluate the classification and segmentation performance of the proposed methods. It is also used to evaluate the tracking performance of the region tracking method proposed herein. The full details on this database are described in Section 2.5 of Chapter 2.
2. **COPD medical dataset.** This is a medical dataset consisting of HRCT lung images of 19 healthy individuals and 13 COPD patients, comprising a total of 32 subjects. The HRCT image volumes obtained from this dataset have a size of $256 \times 256 \times 256$. This data was acquired as a part of a study into the application of imaging to the characterization of the phenotypes of COPD. Written

informed consent was given and signed by all subjects. The study was approved by the Southampton and South West Hampshire Local Research Ethics Committee (LREC number: 09/H0502/91) and the University Hospital Southampton Foundation Trust Research and Development Department. The study was conducted in the Southampton NIHR Respiratory Biomedical Research Unit.

3. **The Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI).** This is a publicly accessible database for medical imaging research (Armato III et al. 2011). The database contains 1,018 clinical chest CT scans, each of which is associated with an XML file that includes detailed information on each case, such as the nodule locations and annotations. This database was used in this research to evaluate the classification performance of the proposed rotation invariant descriptor.
4. **Cilia dataset.** This dataset consists of microscopic images of cilia for the diagnosis of Primary Ciliary Dyskinesia (PCD) with a size of 512 × 512 and 512 frames. This dataset is used to evaluate the texture-based tracking method proposed in this research.

1.5 Publications

The following publications are based on this research:

P1. Almakady, Y., Mahmoodi, S., Conway, J. and Bennett, M. "Volumetric texture analysis based on three-dimensional gaussian markov random fields for COPD detection." In Annual Conference on Medical Image Understanding and Analysis (MIUA), pp. 153-164. Springer, Cham, 2018.

P2. Almakady, Y., Mahmoodi, S. and Bennett, M. "Gaussian Markov Random Fields-Based Features for Volumetric Texture Segmentation." In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 212-215. IEEE, 2019.

P3. Almakady, Y. and Mahmoodi, S. "Texture-Based Region Tracking Using Gaussian Markov Random Fields for Cilia Motion Analysis." In 2019 IEEE International Conference on Image Processing (ICIP), pp. 1292-1296. IEEE, 2019.

P4. Almakady, Yasseen, Sasan Mahmoodi, Joy Conway, and Michael Bennett. "Rotation invariant features based on three dimensional Gaussian Markov random fields for volumetric texture classification." Computer Vision and Image Understanding 194 (2020): 102931.

P5. Almakady, Y., Mahmoodi, S. and Bennett, M., ”Adaptive Volumetric Texture Segmentation based on Gaussian Markov Random Fields Features.” Submitted to Pattern Recognition Letters.

P6. Almakady, Y., Mahmoodi, S. and Bennett, M. ”COPD Detection Using Three-Dimensional Gaussian Markov Random Fields Based on Binary Features.” The 27th IEEE International Conference on Image Processing 2020 (ICIP). [Accepted].

1.6 Thesis overview

The structure of this thesis is outlined below.

Chapter 2. This chapter discusses the importance of volumetric texture and the approaches that are used to analyse features extracted from it. In addition, it presents and compares the neighbourhood configurations used to extract volumetric texture from volumetric images. Moreover, the various applications of volumetric texture analysis are highlighted and examples of methods from each application are provided.

Chapter 3. This chapter begins by explaining random fields, MRFs, GMRFs, and the related analysis methods. The extension of GMRF to 3D GMRF to extract texture features from volumetric images using a spherical neighbourhood configuration with uniformly distributed neighbours is then introduced. A number of implementation issues that arise during parameter estimation for the 3D GMRF model are also discussed. Further, this chapter presents the texture features extracted by 3D GMRFs. These are employed for volumetric texture classification, segmentation, and texture-based region tracking in the following chapters.

Chapter 4 [*Publications* P1, P6]. This chapter introduces the 3D GMRF-based method for volumetric texture classification where new descriptors are constructed based on features extracted by 3D GMRF. This chapter evaluates the proposed method in comparison to other methods in the literature, demonstrating the significance of texture features extracted from the 3D space in providing more discriminative information.

Chapter 5 [*Publications* P4]. In this chapter, the method proposed in Chapter 3 is further improved by extending the descriptor to rotation invariant descriptor with the aim of extracting texture features from volumetric images with any rotated texture patterns. This chapter reviews the existing methods in the literature and how they achieved the rotation invariance. We demonstrate the robustness of our proposed method, including the rotation invariance property and its ability to extract texture features under various random rotations. We also show that the proposed method does not require a large dataset like that required by deep learning-based methods to achieve high classification performance. This finding suggests that the method proposed here is suitable for medical applications where small datasets are common for many reasons discussed further in this chapter.

Chapter 6 [*Publications* P2, P5]. This chapter introduces a new method for volumetric texture segmentation based on 3D GMRF features. A few issues related to the uncertainty principle and the inability of the model to capture different texture patterns are also addressed. The chapter further presents a solution in which an adaptive approach is adopted in the method to segment textures found in volumetric images, overcoming the need to adjust the variables during filtering and parameter estimation. The proposed method is also used to detect the parts of lungs affected by COPD.

Chapter 7 [*Publications* P3]. This chapter addresses the challenges of textured region tracking and proposes a method based on features extracted by 3D GMRF. We demonstrate the ability of this method to track the textured region and extract the trajectory of the ciliary motion, which appear as moving texture in the video image of cilia.

Chapter 8. This final chapter summarises the thesis and proposes a number of directions for future work.

Chapter 2

Volumetric texture

2.1 Introduction

Texture is intrinsic to human vision as it provides visual cues about objects in our natural world. Texture is defined as 'the feel, appearance, or consistency of a surface or a substance' according to Oxford dictionaries.[1] More specifically, Haralick (1979) describes texture images as 'the number and types of its primitives and the spatial organisation or layout of its primitives'. Finding a universal definition of texture is difficult, however, due to the diversity of natural and artificial texture images (Jain et al. 1990). The definition of texture proposed by Haralick (1979) can be slightly extended to describe the 3D texture as an arrangement of 3D texture primitives represented by a sequence of voxels.

When dealing with 3D texture, the underlying texture can be divided into dynamic texture and volumetric texture (Maani et al. 2014). Dynamic texture (also referred to as space-time textures or temporal textures) is the texture represented by the space-time (x, y, t) and it is the accumulation of spatial texture defined in the (x, y) domain over the time t. Examples of dynamic texture include image sequences of 2D texture acquired at different times, such as video of fire, sea waves, wavy trees, etc. Volumetric texture, on the other hand, refers to texture found in the 3D space defined by (x, y, z). This type of texture is commonly found in medical images such as MRI and computed tomography (CT) scans. An additional type of texture that can be considered and referred to as 2.5D texture, which is the texture that exists on the surface of objects (Depeursinge, Foncubierta-Rodriguez et al. 2014).

---

[1] https://en.oxforddictionaries.com/definition/texture, as of 4 May 2020.

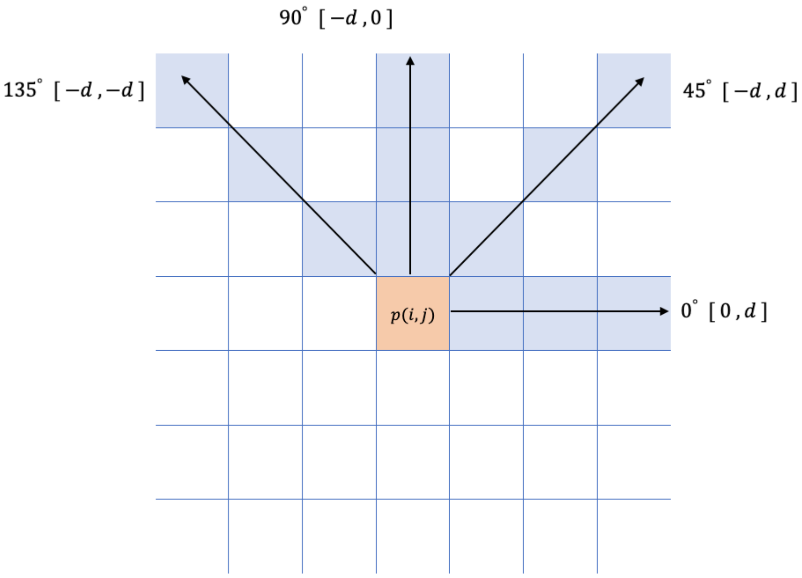


FIGURE 2.1: GLCM with four directions (i.e., $0°, 45°, 90°, 135°$) and d distance.

2.2 Current approaches to volumetric texture feature extraction

The significance of textures as visual cues in the computer vision discipline has led to the development of different methods of texture analysis. Various methods have been proposed for volumetric texture analysis. However, most of these methods are based on 2D approaches that deal with volumetric texture as 2D slices, which could ignore rich information contained within the volumetric data. Methods of volumetric texture feature extraction generally aim to extract features and employ them as an optimal discriminative descriptor of textures, which can consequently be used for several tasks, including texture classification, segmentation and synthesis. Approaches to texture feature extraction can be broadly categorised into statistical-, structural-, spectral- and model-based methods (Tuceryan et al. 1993). However, this categorisation does not include deep learning approaches, which have received much attention recently. Therefore, 3D approaches to volumetric texture feature extraction can follow the categorisation proposed by Tuceryan et al. (1993) in addition to deep learning methods.

2.2.1 Statistical-based methods

In statistical approaches, texture is described by using statistical features. An example of a statistical-based method is the well-known grey-level co-occurrence matrices (GLCM) method proposed by Haralick et al. (1973). It is among the earliest attempts to quantify the co-occurrence of values of pixels separated by a given distance and in a specific direction (e.g., Figure 2.1). Texture descriptors, therefore, can be constructed from each

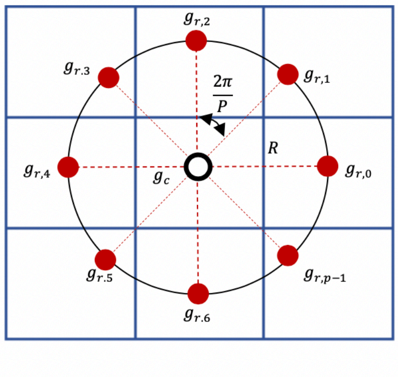


Figure 2.2: Visualisations of LBP descriptor.

matrix by calculating a particular textural feature. Haralick et al. (1973) propose a set of textural features that can be extracted to form spatial dependence matrices of grey-level values, including angular second moment, contrast, correlation, variance, inverse difference moment, and entropy. This method is extended to 3D-GLCM for volumetric texture analysis and is used for volumetric texture classification, segmentation and characterisation (Quellec et al. 2010; Mahmoud-Ghoneim et al. 2003; Kovalev, Petrou et al. 1999; Kurani et al. 2004; Pena et al. 2018; Othmen et al. 2013).

Another method based on statistical approaches is the run-length, which was first introduced by Galloway (1974). A grey-level run is defined as a set of consecutive collinear voxels with the same grey-level value. Once the run length matrices is computed, common statistics such as short run emphasis and long run emphasis are extracted from these matrices for texture analysis. Run-length extension to 3D is proposed to characterise volumetric texture found in CT scans of the chest and abdomen (D.-H. Xu et al. 2004; Jing Zhang et al. 2012; De Nunzio et al. 2011; Basu et al. 2011).

2.2.2 Structural-based methods

Structural-based methods aim to characterise texture by relying on texture primitives and the spatial arrangement of these primitives. Among the structural methods is the local binary pattern (LBP; Ojala, Pietikäinen and Harwood 1996), which has received considerable attention due to its simplicity and effectiveness for texture analysis. LBP originally worked in the 3-by-3 neighbourhood and was extended later by Ojala, Pietikäinen and Maenpaa (2002) to further improve ways of dealing with multiresolution and rotated textures. Specifically, $\text{LBP}_{P,R}$ is obtained by thresholding the equally spaced neighbouring points P arranged on a circle of radius R against the central pixel in the local neighbourhood, where all neighbours have values higher than or equal to the central pixel are assigned to a value of one, and the remainder to a value of zero. The associated values of those neighbouring points are then read sequentially in a clockwise manner, producing a binary number. The LBP of a given pixel is obtained by,

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \quad s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{2.1}$$

where P is the number of neighbours surrounding the central pixel g_c and g_p is the neighbour pixel value. Figure 2.2 illustrates the circularly symmetric neighbourhood used by LBP. The distributions of LBP codes are then computed to describe the local texture. This method is extended to extract texture features from 3D images and is used to solve various problems, such as volumetric and dynamic texture classification (Paulhac, Makris and Ramel 2008; Mishra et al. 2011; S. Liu et al. 2011; Päivärinta et al. 2011; Fehr and Burkhardt 2008; Morgado et al. 2013; X. Zhao et al. 2019; Citraro et al. 2017; G. Zhao et al. 2007).

Extraction of texton features from the images is another structural-based method. It is introduced by Julesz (1981) as primary elements of pre-attentive texture perception. The concept of texton is adopted by Leung et al. (2001) to build textons from textured images. This is performed by filtering the textured image with a bank of filters producing a filter response vector at each pixel corresponding to different lighting and viewing directions. The resulting filter responses are then clustered using k-mean, where mean vectors known as textons are used to represent the dominant features in the image. The work presented by L. Wang et al. (2017) extends the idea of texton to extract 3D texture features from MRI images for prostate cancer detection.

2.2.3 Model-based methods

In this approach, texture is explored by constructing models to represent volumetric images, in which the model estimated parameters are used as texture features. A popular method for modelling images is Markov random fields (MRF) and its subclass GMRF, which can capture the essential structures of texture, and it is emerging as a popular choice for modelling images (Tuceryan et al. 1993). MRF and GMRF extension to 3D are introduced for volumetric texture analysis (Shafer et al. 2011; Ranguelova et al. 1999).

Modelling texture with a fractal model is another choice for texture feature extraction based on model approaches. The fractal model is used to measure an object via self-similarity under different scales and resolutions (Al-Kadi 2017). The essential characteristic of the fractal is the fractal dimension, which measures the object's dimension at different scales (Al-Kadi 2017; Casanova et al. 2016) and can be computed using different methods, including the Hausdorff–Besicovitch dimension and the Bouligand–Minkowski dimension (Florindo et al. 2012; Al-Kadi 2017). The fractal dimension with the same value at different scales indicates the homogeneity of the texture, whereas in a multifractal case the dimension changes for different scales (Conci et al. 2000). The

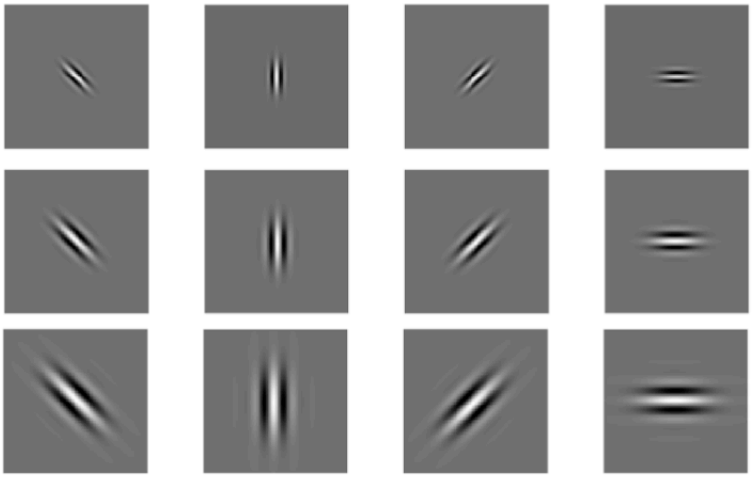


Figure 2.3: Gabor filters.

fractal model plays a vital role in biomedical texture analysis, such as the application of tumour tissue characterisation (Al-Kadi 2017). In the method presented by Lopes et al. (2011), an extension of the fractal model to analyse 3D texture in volumetric images is proposed for volumetric texture characterisation. This method utilises the fractal geometry with fractal and multifractal features to detect tumours in MRI images of the brain.

2.2.4 Spectral-based methods

In this approach, the volumetric image is convolved with a set of filters to extract texture features. Various methods are proposed for volumetric texture analysis by constructing a bank of filters that are able to extract different measurements and use them as texture features. An example of these methods is Gabor filters, in which an image is convolved with a bank of filters with different scales and orientations (e.g., Figure 2.3) , and the produced response filter can be employed as a texture feature. The Gabor filters method is extended by Gonçalves et al. (2012) to extract 3D texture features and used for dynamic texture recognition. Moreover, the method introduced by Almaev et al. (2013) utilises Gabor filters to extract texture features from dynamic texture based on the orthogonal planes technique. The extended 3D Gabor filters method is also employed by Fernández et al. (2000) for 3D seismic segmentation.

Wavelet-based feature extraction is a spectral-based method allowing the characterisation of texture with different scales and directions. Various methods are proposed to extract wavelet-based features from volumetric images. Wavelet-based is extended to 3D in which 3D-Riesz wavelets are employed to extract texture features from the volumetric images and utilise them for different tasks, including medical image analysis (Cid et al.

2017; Cirujeda, H. Müller et al. 2015; Cirujeda, Cid et al. 2016; Depeursinge, Pad et al. 2015; Barra et al. 2000).

2.2.5 Deep learning-based methods

Recently, deep learning approaches such as convolutional neural networks (CNNs) have made remarkable contributions to 3D image analysis, including image classification. The primary difference between the above hand-crafted texture feature approaches and deep learning approaches is that features in deep learning are learned adaptively and automatically through the training process, whereas in hand-crafted methods features are designed manually.

Texture analysis based on these methods can be performed by combining texture features with deep learning extracted features. Gabor filters, for example, are used to generated Gabor features that are then fed to a CNN to reduce the training time and complexity of CNN (Sarwar et al. 2017; Yao et al. 2016). Gabor filters are incorporated into CNN to provide robust features against various image transformations, such as transitions, scale and rotations (Luan et al. 2018). To classify lung nodules in chest CT scans, the GLCM texture feature is fused with CNN features extracted from the slices of the volumetric data (Xie, Jianpeng Zhang, Xia et al. 2018).

Deep learning methods are also utilised for dynamic texture analysis. Texture is analysed by CNN for synthesising dynamic textures (Funke et al. 2017), classifying dynamic textures (Andrearczyk and Whelan 2018) and 3D mask face anti-spoofing (Shao et al. 2017).

2.3 Volumetric texture neighbourhood configurations

2.3.1 Neighbourhood configurations

Texture features are commonly extracted in a small local neighbourhood with a specific displacement and given direction, such as in methods of LBP, GMRF and GLCM. In volumetric images, texture is extracted by using different configurations to select the neighbour voxels around the central voxels. These configurations can be as follows:

- *3D neighbourhood*

  In the 3D neighbourhood, voxels are distributed in 3D space, such as on the surface of a sphere or a cube centred at the central voxels (e.g., Figure 2.4a). Selecting the optimal number of neighbouring voxels depends on the method and application. Generally, selecting a small number of voxels on the surface of a

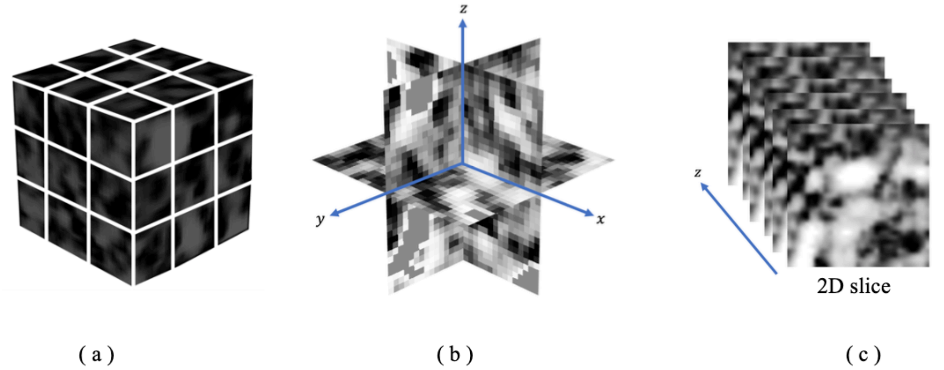


Figure 2.4: Overview of neighbourhood configurations. (a) 3D neighbourhood. (b) Orthogonal planes. (c) Sequences of 2D slices.

sphere or a cube requires less computation time but ignores some details of texture. Conversely, although many neighbouring voxels capture fine details of texture, they capture noise as well as this requires more computation time. Methods presented in (Citraro et al. 2017; Paulhac, Makris and Ramel 2008; Fehr and Burkhardt 2008; Fehr 2007; Banerjee et al. 2012) adopt the spherical neighbourhood with different sampling rates.

- *Orthogonal planes*

  Methods that adopt the orthogonal planes-based configuration aim at sampling points in three 2D slices or planes intersecting at the central voxel to extract texture features from 3D space (e.g., Figure 2.4b). The features extracted from each plane can then be processed separately and concatenated depending on the method. The LBP-TOP methods proposed by G. Zhao et al. (2007) and Ciompi et al. (2015) extract LBP codes from each 2D slice arranged orthogonally and concatenate them to produce the feature vector.

- *Sequences of 2D slices*

  Extracting features from 3D data is commonly performed using sequences of 2D slices (e.g., Figure 2.4c). It allows exploitation of methods developed for 2D images to be used for volumetric images in a slice by slice manner. The disadvantage associated with this technique is that it prevents exploration of the rich details provided by the volumetric data. The VLBP method proposed by G. Zhao et al. (2007) adopt a sequence of 2D slices to extract texture features.

2.3.2 Comparison

Volumetric texture can generally be analysed using 2D or 3D approaches. Many studies have found that texture features extracted from volumetric data have more discriminative power than features extracted from 2D slices. This aligns with the findings presented by Han et al. (2015), where Haralick texture features are extracted using both 2D and 3D strategies and it is found that better classification performance is achieved by considering the 3D data than when a 2D slice is used. Similarly, three strategies are used by Yan et al. (2016) to process the volumetric image of a medical dataset of a lung nodule: slice-level 2D, nodule-level 2D, and nodule-level 3D using CNN. The results show better performance when features are extracted from 3D space. Likewise, a comparison is carried out by Paulhac, Makris and Ramel (2008) between 2D and 3D LBP methods, demonstrating that 3D LBP outperforms 2D LBP in different datasets involved in the experiment. Employing the 2D strategy can be beneficial when a deep learning method is used, as it helps to provide more training data by training each single 2D slice extracted from specific volumetric data (Xie, Jianpeng Zhang and Xia 2019).

2.4 Applications of volumetric texture analysis

The fundamental property of volumetric data is that it contains rich information which can be analysed for various applications. Volumetric texture analysis has several practical applications, which include the following:

2.4.1 Medical image analysis

Volumetric texture analysis has been employed extensively in medical imaging and diagnostic radiology. Many diseases have symptoms that appear as textural changes in the medical images of soft tissue. However, these changes in the texture of images are difficult to analyse by human perceptions, especially for some diseases such as cancer, for which normal cells are very similar to abnormal cells, leading to inaccurate diagnoses (Faust et al. 2018). Therefore, volumetric texture-based methods are essential for medical image analysis to provide reliable, fast and accurate diagnoses. Volumetric images acquired by different image modalities show that they are composed of texture starting from cell level to the organ level (Depeursinge, Foncubierta-Rodriguez et al. 2014). Consequently, texture characteristics are employed to automatically diagnose a wide range of diseases.

Texture patterns of a lung affected by COPD, for instance, are different from those appearing in a normal lung, which is utilised for COPD detection (Ye Xu et al. 2005; J. Yang et al. 2017; Mariolis et al. 2010; Korfiatis et al. 2009). Moreover, volumetric texture analysis is employed for brain disease diagnosis, such as Alzheimer's disease classification (Hett et al. 2018; X. Li et al. 2011) dyslexia diagnosis (El-Baz et al. 2008), and epileptogenic lesion quantification (Toro et al. 2013). Moreover, it can be further used for breast cancer detection (Waugh et al. 2016; Jalalian et al. 2017; Thibault et al. 2017), liver lesions diagnosis (Upadhyay et al. 2012; Danciu et al. 2012), and prostatic adenocarcinoma segmentation (Madabhushi et al. 2003).

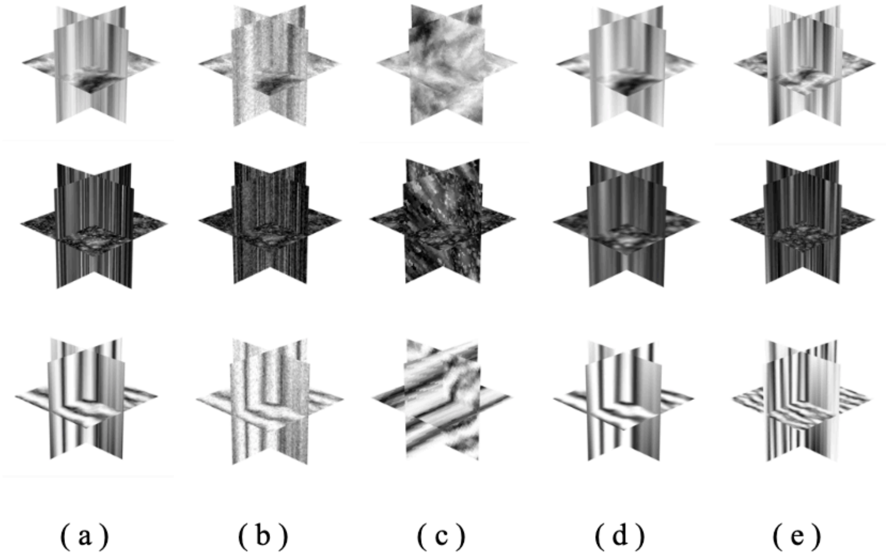


Figure 2.5: Examples of the Interpolated dataset from the RFAI database with various transformations. (a) Normal. (b) Noise. (c) Rotate. (d) Smooth. (e) Subsampling.

2.4.2 Seismic texture analysis

Volumetric texture analysis has also found its applications in geology and the petroleum industry. The volumetric data is acquired by transmitting sound waves into the Earth's subsurface, generating seismic reflection volume from the returns of these waves (Patel et al. 2010). Density changes between rocks reflect the sound waves back to the surface, which are visualised as volumetric data known as 3D seismic data (Chopra et al. 2006; Yenugu et al. 2010; Mondol 2010). Texture is a key attribute of 3D seismic data that can be analysed to interpret information represented in the data. Texture-based methods such as GLCM are used to analyse the texture attribute of volumetric surface seismic data (Chopra et al. 2006; Yenugu et al. 2010). Moreover, the voxel co-occurrence matrix (VCM) is employed by D. Gao (2003) for seismic texture analysis, allowing evaluation of textural features along different directions in 3D space. The employment of texture features to analyse 3D seismic data demonstrates a better understanding compared to the other attributes of this data (Chopra et al. 2006; Yenugu et al. 2010).

2.5 Volumetric texture datasets

There are few synthetic volumetric textures datasets available for method evaluations. The RFAI database[2], which is created by Paulhac, Makris, Ramel et al. (2009) for evaluating texture classification and segmentation methods, is one of the available databases

---

[2] http://www.rfai.li.univ-tours.fr/PublicData/3D_Textures/3Dsynthetic_images_database.html, as of 15 April 2020.

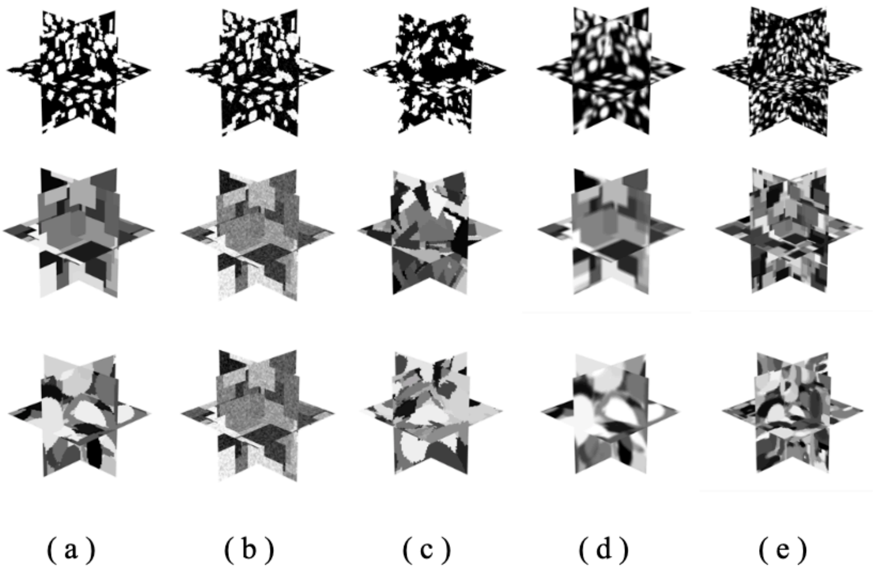


Figure 2.6: Examples of the Geometric dataset from the RFAI database with various transformations. (a) Normal. (b) Noise. (c) Rotate. (d) Smooth. (e) Subsampling.

Table 2.1: Summary of RFAI database used for classification

Dataset	No. of Classes	Samples per class	Total samples per subset	Total samples
Interpolated	30	10	299	1495
Geometric	25	10	250	1250
Fourier	15	10	150	750
Mixed texture	25	10	250	1250

for synthetic textured 3D images. This dataset is constructed using a 2D dataset such as Brodatz and fractal textures. Four different methods are employed to generate this dataset. The first method interpolates two or more different 2D texture images to construct a volumetric texture where the 2D textures used in this method are not identical (e.g., Figure 2.5). The second method uses geometric shapes such as spheres and cubes to generate volumetric textures (e.g., Figure 2.6). The third method employs the Fourier transformation to synthesise volumetric textures (e.g., Figure 2.7), whereas the fourth method combines the three previous methods to generate volumetric textures (e.g., Figure 2.8). Table 2.1 lists the number of classes contained within each dataset. These methods produce four different categories, with each category containing five types of textures according to the applied distortion: normal, rotated, Gaussian blur, Gaussian noise, and subsampling. Each of these types contains several classes and each class comprises ten volumetric images with a size of $64 \times 64 \times 64$. The database also offers a dataset with three categories for segmentation problems. These categories of images containing two, three, and four texture classes, respectively, and each category contains five texture images, for a total of 15 images each with dimensions $128 \times 128 \times 128$. This diversity of synthetic textures provides an appropriate way of evaluating texture analysis methods under different conditions and challenges.

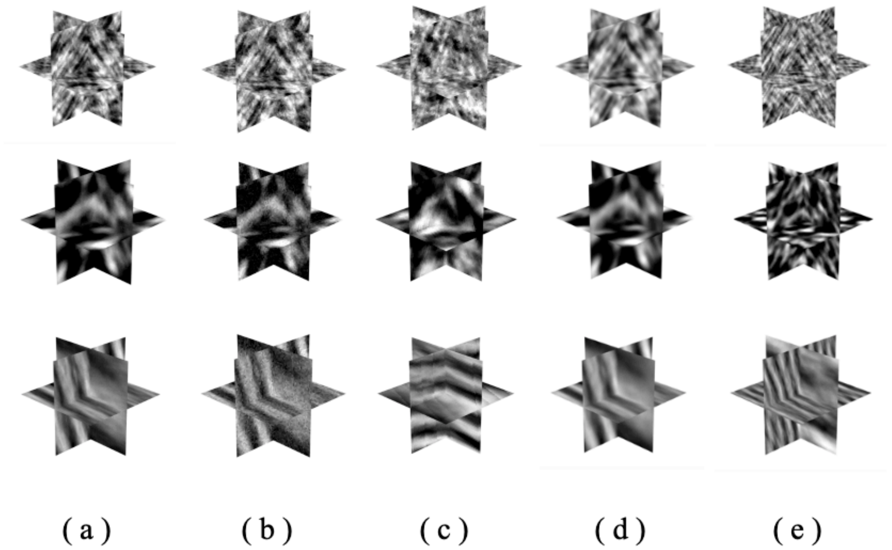


Figure 2.7: Examples of the Fourier dataset from the RFAI database with various transformations. (a) Normal. (b) Noise. (c) Rotate. (d) Smooth. (e) Subsampling.

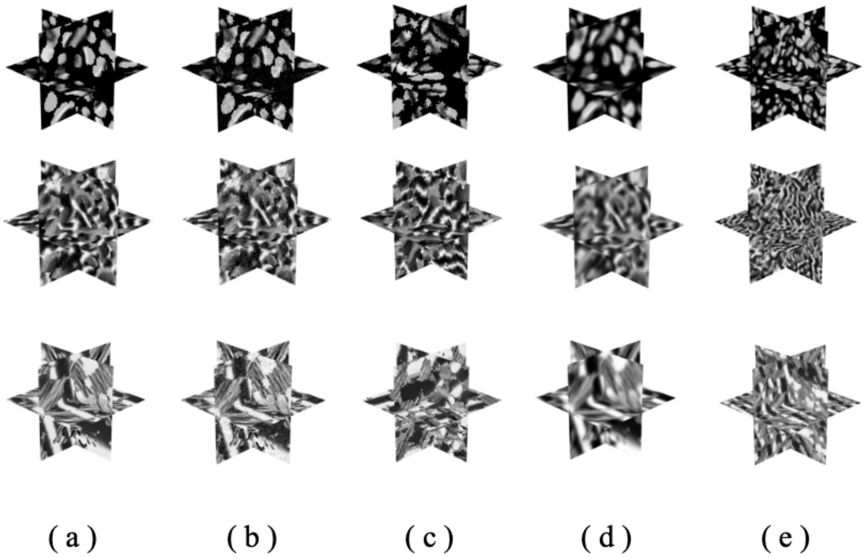


Figure 2.8: Examples of a dataset from the RFAI database generated by a mixture of the three previous methods (Fourier, Interpolation and Geometric methods) with various transformations. (a) Normal. (b) Noise. (c) Rotate. (d) Smooth. (e) Subsampling.

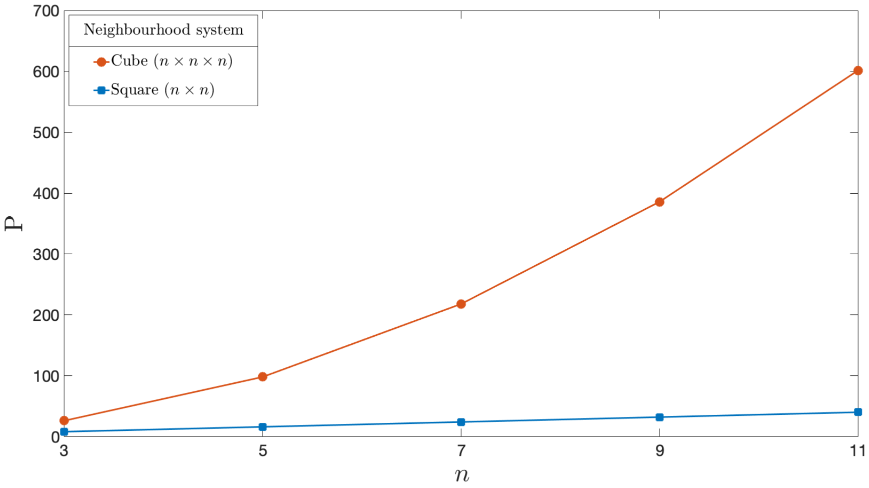


Figure 2.9

2.6 Challenges

There are several issues and challenges associated with analysing texture in volumetric images. These issues and challenges make the extension of 2D texture analysis methods into 3D a complex task. The sampling points of neighbours is a key issue resulting in consequential challenges. A basic neighbourhood scheme in 2D, that is, a square neighbourhood, has a less number of sampling points compared with a cubic neighbourhood in 3D. For example, an $n \times n$ square neighbourhood, where $n = \{3, 5, 7\}$, results in $\{8, 16, 24\}$ neighbours surrounding the central pixel. However, for an $n \times n \times n$ cube neighbourhood, the number of neighbours surrounding the central voxel increases significantly to $\{26, 98, 218\}$. Figure 2.9 indicates that the number of the required neighbours increases significantly in 3D images compared to 2D images for various sizes of neighbourhood.

This extreme increase in sampled points compared to 2D is due to the involvement of the third dimension, which also raises another issue regarding the rotation invariance. In 2D texture, the order of points that are sampled on a circle, which is essential for the achievement of rotation invariance, is maintainable. This is because the rotation of texture results in the rotation of the neighbour sets either clockwise or anticlockwise, and therefore it is easy to keep the order of points that are rotated only in two orientations. However, in 3D texture with a spherical neighbourhood scheme, the order of neighbourhood voxels is lost when texture is rotated due to the increase in dimensionality (Banerjee et al. 2012), and it becomes a complicated task to achieve rotation invariance depending on the points order notion.

The dimension of the feature vector is another issue that is associated with the growth of neighbouring voxels. The length of the resulting histogram that describes the 3D texture is exponentially increased in accordance with the growth of sampling voxels (S. Liu et al. 2011). This leads to extensive computation time and could cause instability during feature matching due to the large number of histogram bins. Therefore, the construction of an optimal descriptor to describe the volumetric texture requires an efficient method capable of dealing with such challenges.

2.7 Conclusion

This chapter presents an overview of volumetric texture with current approaches in the literature regarding volumetric texture feature extraction. The current approaches presented in this chapter are grouped into statistical-, structural-, spectral-, model- and deep learning-based methods with highlighting their extension into 3D methods. In addition, this chapter describes the neighbourhood configurations used to extract texture features from volumetric images with a comparison provided between 2D and 3D neighbourhood schemes. Subsequently, the applications of volumetric texture analysis with examples of some methods proposed in each application area is presented. The available volumetric texture dataset is described in this chapter, which also elaborates the methods used to generate the dataset as well as its structure. Finally, the challenges arising when dealing with volumetric images are discussed, emphasising the need to develop a method that is capable of overcoming such challenges.

Chapter 3

Three-dimensional Gaussian Markov random fields

3.1 Introduction

In texture analysis based on model approaches, models are generated to describe the underlying texture features of a given image. Two of the most popular models for texture analysis are Markov random fields (MRFs) and their sub-class Gaussian Markov random fields (GMRFs). Such models are useful in modelling the spatial dependencies between a pixel and its neighbours (Y. Zhao et al. 2007). This chapter investigates the extension of GMRFs into three dimensions (3D GMRFs) to extract texture features from volumetric images.

This chapter starts by explaining random fields, MRFs, GMRFs, and the related analysis methods in Sections 3.2 and 3.3. Section 3.4 introduces the extension of the GMRF to the 3D GMRF to extract texture features from volumetric images using a spherical neighbourhood configuration with uniformly distributed neighbours. In this section, a number of implementation issues that arise during parameter estimation for the 3D GMRF model are also discussed. Further, Section 3.5 of this chapter presents the texture features extracted by the 3D GMRF. These are employed for volumetric texture classification, segmentation, and texture-based region tracking in the following chapters. Lastly, Section 3.6 summarises the chapter.

3.2 Random fields

In image analysis based on random fields, images can be represented as a two-dimensional (2D) random field which is a set of random variables arranged in a 2D regular lattice (R. Chellappa 1989; Besag 1974). The regular lattice consists of a set of sites, each site

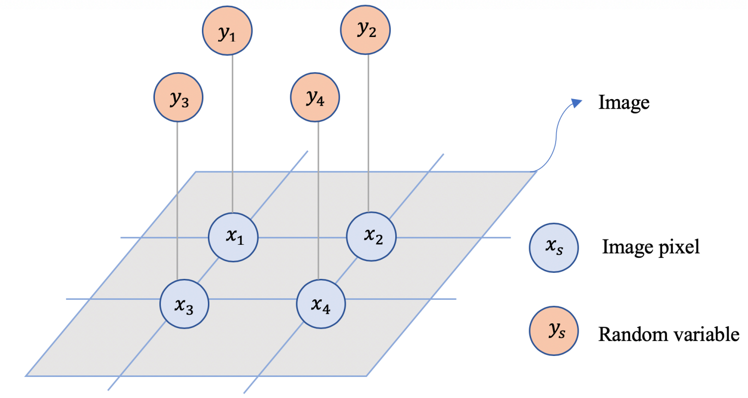


FIGURE 3.1: Graphical representation of a 2D regular site.

representing a pixel and being associated with a random variable. For example, given a 2D image of size $n \times n$, a regular lattice S whose sites are at locations represented by $s = \{i, j\}$ is defined for the image as $S = \{s = (i, j) \mid 1 \leq i, j \leq n\}$. If y_s are the associated random variables, each with site s, then the random field is defined as $Y = \{y_s | s \in S\}$. The spatial dependence of the random field on its neighbours is expressed locally by the conditional probability model $p(y_s | y_t, t \neq s, t \in S)$. When texture analysis is being considered, an assumption is made that the random field is conditionally independent of all random fields except its neighbours, which implies the Markovian property and thus the name "Markov random fields."

3.3 Markov random fields

A Markov random field is defined as a probabilistic model for random variables over a group of nodes that are connected to one another in an undirected graph (Blake et al. 2011; Witten et al. 2017). The random fields defined by an MRF satisfy the Markovian property, according to which a node is conditionally independent from all other nodes except its neighbours (Blake et al. 2011; Rue et al. 2005; Petrou et al. 2006). In terms of images, each node in the graph corresponds to a pixel arranged in a specific structure. MRFs structure illustrated in Figure 3.1 is the most popular choice for representing image pixels (Witten et al. 2017). The direct dependency of a random field on its neighbours provides an appropriate way to model textures, as the pixels underlying the textures also have local dependencies on their neighbours (S. Z. Li 2009). Therefore, MRFs have been widely used for modelling vision problems, including images texture analysis (C. Wang et al. 2013).

In computer vision and image processing, MRFs models can be used for both the low-level and high-level processing of images. MRFs models are more popular for low-level

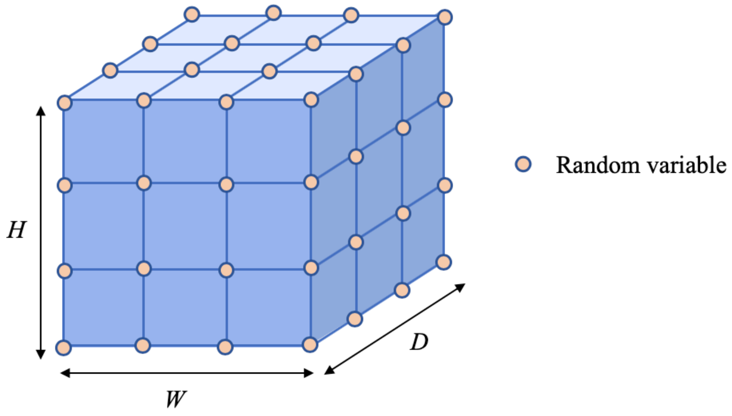


FIGURE 3.2: Illustration of a 3D regular site.

processing applications, including image segmentation, edge detection, and texture analysis. High-level processing is used in object matching and recognition (S. Z. Li 2009). There are two approaches to representing MRFs: globally in terms of joint probability or locally in terms of conditional probability. Generally, the latter is preferable due to its simplicity and computational efficiency (S. Z. Li 2009).

Local conditional probability distributions that obey Gaussian functions are called Gaussian MRFs (C. Wang et al. 2013; Rue et al. 2005). GMRFs can be used for modelling image textures by employing the model parameters as texture features, which can then be used for texture analysis. Altogether, GMRFs have been widely used for image texture classification. To begin with, Rama Chellappa et al. (1985) and Kashyap et al. (1982) introduce the employment of GMRFs for texture classification using the estimates of the model parameters as features. Later, Cohen et al. (1991) improve these GMRFs by incorporating rotation and scale invariance properties into a GMRF model and using it for texture classification. In addition, Y. Zhao et al. (2007) investigate the neighbourhood techniques used in GMRF-based methods and found that all neighbouring pixels could be treated equally when a GMRF is generated. Accordingly, a GMRF model is proposed in which the neighboring pixels are regarded according to their distances from the central pixel. Further, Hafner et al. (2009) combine GMRFs with wavelet transforms by performing parameter estimation of a GMRF on the wavelet domain, rather than directly on the image, to extract features for image classification.

Recently, GMRFs have been improved to characterize local textures in small regions (Dharmagunawardhana et al. 2016) . According to this method, rotation-invariant texture features are extracted from local patches using estimation windows for texture classification. Similarly, Deng et al. (2004) propose an anisotropic circular GMRF (ACGMRF) model to extract rotation-invariant features that can be employed for texture classification.

More recently, MRFs have been incorporated into deep-learning-based methods to post-process classification maps. Further, Danilla et al. (2017) propose a method that combines convolutional neural networks (CNNs) with MRFs, where the output of a CNN is used as an input for an MRF model to produce a classification map. Likewise, Cao et al. (2018) propose a method that integrates spectral and spatial information using the combination of a CNN and an MRF.

3.4 Three-dimensional Gaussian Markov random fields

Modelling of an image using a GMRF is normally achieved by defining a set of sites Ω corresponding to the voxels in the image v. To extend a GMRF to extract features from volumetric images, the regular lattice site is defined as a set of points arranged in a 3D space as shown in Figure 3.2. Let $\Omega = \{v = (i, j, k) \mid 1 \leq i \leq H, 1 \leq j \leq W, 1 \leq k \leq D\}$ denote a set of points indexed by (i, j, k) on an $H \times W \times D$ 3D lattice corresponding to voxels in a 3D image volume. The local conditional probability density function of the intensity value g_v at location v is defined by:

$$p(g_v|y_{v+r}, r \in V_v) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}\left(g_v - \lambda - \sum_{r\in V_v} \alpha_r(y_{v+r} - \lambda)\right)^2\right\} \tag{3.1}$$

where V_v represents the neighbours of the voxel at location v, λ is the mean intensity of the neighbourhood patch, and α_r is a set of interaction coefficients that represent the influence on a voxel with intensity value y_{v+r} located at relative position r (Dharmagunawardhana et al. 2016; Petrou et al. 2006). The neighbourhood scheme adopted here is sampled voxels over a sphere surface with radius R, so that $V_v \in \{\theta, \phi \mid 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi\}$, where θ and ϕ represent the angles of colatitude and longitude respectively on the spherical neighbourhood using spherical coordinates to sample the neighbouring voxels, and $|V_v|$ is equivalent to the number of voxels surrounding the central voxel. Figure 3.3 presents the sampled voxels over a sphere equivalent to the neighbours y_r, where the center of the sphere corresponds to the grey scale of the central voxel g_v. The 3D GMRF model parameters α_r and σ^2 in Eq. 3.1 are unknown and need to be estimated by finding a best fit of model of 3.1 to the texture data.

3.4.1 Parameter estimation

The probability distribution function presented in Eq. 3.1 is fitted to texture data by estimating the unknown parameters α_r, σ^2, and λ. Two common methods can be used to estimate the parameters: the maximum likelihood estimation (MLE) (X. Gao et al. 2010) or the least square estimation (LSE) (Sorensen, Nielsen et al. 2012) . Although both methods produce adequate results and lead to the same set of equations in terms of

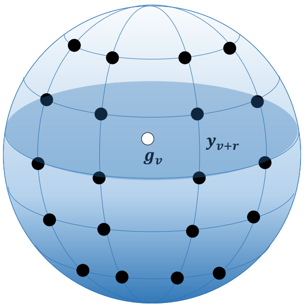


FIGURE 3.3: Graphical representation of voxels y_{v+r} sampled over a sphere surface where g_v is the central voxel.

the GMRF, MLE has many advantages over LSE (Genschel et al. 2010). Aside from being fast and easy to implement, it provides an optimal estimation in large samples while also being consistent in small samples. Maximum a posteriori (MAP) estimation is also a technique that can be used to estimate the model parameters by incorporating prior knowledge about the parameters. Prior knowledge can include, for instance, texture colour, texture scale or locations of a specific pattern such as in some medical images where specific characteristics of the disease are observed in particular areas. When prior knowledge is available, MAP is better to estimate the parameters. However, our approach is not restricted by assumptions about the texture, and no specific prior information is assumed about the parameters. Hence MLE is preferably selected for use in GMRF model parameter estimation due to its simple implementation. MLE is performed by taking the partial derivative of a log-likelihood function with respect to α, σ^2, and λ and setting it to zero, which leads the following simultaneous equations:

$$Z\left(1-\boldsymbol{\alpha}_v\right)\lambda = \sum_{v\in\Omega_v}\left(g_v-\boldsymbol{\alpha}_v\boldsymbol{y}_v\right) \tag{3.2}$$

$$-\lambda\sum_{v\in\Omega_v} y_k + \sum_{v\in\Omega_v} y_k\boldsymbol{\alpha}_v\left(\boldsymbol{y}_v-\lambda\right) = \sum_{v\in\Omega_v} g_v y_k,\ k\in r \tag{3.3}$$

$$\sigma^2 = \sum_{v\in\Omega_v}\Big(g_v-\lambda-\boldsymbol{\alpha}_v\left(\boldsymbol{y}_v-\lambda\right)\Big)^2 \tag{3.4}$$

Here, $Z = H\times W\times D$, $\boldsymbol{\alpha}_v = col[\alpha_r]$, and $\boldsymbol{y}_v = col[y_r]$ for $r\in V_v$. The model parameters α_r, σ^2, and λ are calculated by solving Eqs. 3.2 - 3.4 and are then employed as local features, i.e. $\boldsymbol{f}_v = [{\boldsymbol{\alpha}_v}^T, \sigma^2, \lambda]$, for each voxel v in the image volume.

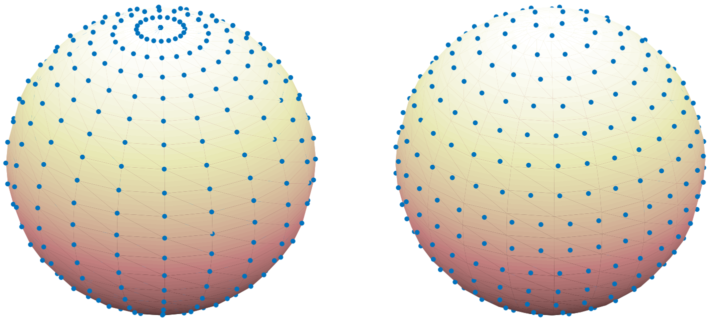


FIGURE 3.4: Illustration of different sampling systems: On the left, points are sampled on the surface of the sphere to result in unequally spaced points. On the right, the points are uniformly distributed on the surface of the sphere.

3.4.2 Local estimation

For a 2D texture image, GMRF model parameter estimation is carried out by sliding an $n \times n$ window over the image to collect samples. However, extracting the features from small regions (patches) and in this way characterising local texture is the preferred method (Y. Zhao et al. 2007; G. Zhao et al. 2007; Rellier et al. 2004; X. Liu et al. 2003; Dharmagunawardhana et al. 2016). We thus adopt the idea of extracting features from 3D patches, according to which local parameter estimations are carried out based on small patches. To collect samples from volumetric images, we propose using a cube with size $w \times w \times w$ for local parameter estimation, (known as an estimation cube; Ω_v). The local estimation is performed by first sliding the estimation cube Ω_v so that it is centred at a voxel v. The neighbourhood sphere is then slid inside Ω_v to collect overlapping samples (i.e., the neighbouring voxels y_r), which are used for parameter estimation. Once the parameter estimation has been performed for the current voxel, Ω_v is moved to the next voxel and the procedure is repeated. As stated before, the neighbourhood scheme adopted in this work is for voxels sampled over the surface of a sphere. To ensure that the voxels on the surface of the sphere are uniformly distributed, an equidistant spherical sampling procedure is used.

3.4.3 Equidistant spherical sampling

In the 2D texture methods proposed by Ojala, Pietikäinen and Maenpaa (2002), equidistant sampling is achieved by the distribution of P equally spaced points, indexed by p, on the circumference of a circle with radius R. The coordinates of such points can be given by $(-Rsin(2\pi p/P), Rcos(2\pi p/P))$. According to the same principle, ensuring the uniformly distribution of points on a sphere is essential to retain their indices. The solution proposed by Deserno (2004) is adopted in this study, according to which equidistant points corresponding to the neighbours of g_v are evenly distributed on the surface of a sphere. The circles of latitude are defined at constant intervals d_θ, while the points on the circles are sampled at constant intervals d_ϕ such that $d_\theta \simeq d_\phi$.

$$d_\theta d_\phi = 4\pi r^2/N \tag{3.5}$$

where r is the radius of a sphere, and N is the number of points required for uniform distribution on the surface of the sphere. Figure 3.4 (right) shows the arrangement of points in an equidistant sampling scheme, resulting in a uniform distribution of points over the sphere surface. In contrast, the most popular sampling scheme samples points on ϕ and θ at equal distances. However, the conversion to the Cartesian coordinate system in this method results in points being bunched around the $(\theta = \pi, \theta = 0)$ and being sparsely distributed around $(\theta = \pi/2)$. Thus, this sampling scheme leads to an incorrect distribution of points (e.g., Figure 3.4 (left)).

3.4.4 Implementation issues

The first issue that arises in the implementation of the present method is with respect to solving Eqs. 3.2 - 3.4, which are nonlinear with respect to λ. Such nonlinear equations are numerically expensive to solve and may also suffer from instability; in other words, solving Eqs. 3.2 - 3.4 implies dealing with nonlinear systems. To avoid nonlinearity, the mean λ of the neighbourhood patch Ω_v is estimated separately by calculating λ and subtracting it from the observation space Ω_v. It is then included as a feature together with the model parameters, the final solution of which leads to:

$$\boldsymbol{\alpha}_v = \left(\sum_{v \in \Omega_v} \boldsymbol{y}_v \boldsymbol{y}_v^T \right)^{-1} \left(\sum_{v \in \Omega_v} \boldsymbol{y}_v g_v \right) \tag{3.6}$$

$$\sigma_v^2 = \frac{1}{|\Omega_v|} \sum_{v \in \Omega_v} (g_v - \boldsymbol{\alpha}_v^T \boldsymbol{y}_v)^2 \tag{3.7}$$

Here, λ is the mean intensity of the neighbourhood patch; it is computed separately for each neighbourhood patch and subtracted from the intensity of the patch to avoid dealing with non-linear equations and their associated issues. The parameters α_r and σ^2 are then calculated based on patches whose means have been subtracted beforehand leading to a set of linear equations.

Another issue is encountered during the parameter estimation stage, when inverting the parameter matrix causes a non-invertibility problem. In particular, the non-invertibility problem arises when the observation matrix $\boldsymbol{y}_v \boldsymbol{y}_v^T$ (i.e, a set of neighbours surrounding g_v) has no inverse and thus becomes a singular matrix with no solution, leading to an inconsistent model. To overcome this issue, we must ensure that the observation matrix is non-singular and invertible. This can be achieved using various methods. One possible direct numerical solution is to use ridge regression by adding a regularization to the observation matrix in a way that makes it non-singular and invertible (Friedman et al. 2001; Björkström 2001). The resulting interaction parameters $\boldsymbol{\alpha}_v$ are then given by:

$$\boldsymbol{\alpha}_v = \left(\sum_{v \in \Omega_v} \boldsymbol{y}_v \boldsymbol{y}_v^T + c^2 I \right)^{-1} \left(\sum_{v \in \Omega_v} \boldsymbol{y}_v g_v \right) \tag{3.8}$$

Here, I is an identity matrix with a size equal to that of the observation matrix $\boldsymbol{y}_v \boldsymbol{y}_v^T$, and c^2 is a regularization parameter. The value c^2 is added to the diagonal of the observation matrix $\boldsymbol{y}_v \boldsymbol{y}_v^T$ before the inversion operation to make it non-singular. The selection of a value for c is not straightforward; it is an empirically determined value chosen to maximize the accuracy of the classification.

3.5 Texture features

Local estimation of the 3D GMRF model parameters leads to a set of parameters that have been extracted from each individual voxel in the volumetric image, and this forms a feature vector that can be represented as follows:

$$\boldsymbol{f}_v = [\boldsymbol{\alpha}_v{}^T, \sigma_v^2, \lambda_v] \tag{3.9}$$

The extraction of the feature vector $\boldsymbol{f}_v$ at each voxel in the volumetric image produces a set of 3D parameter maps corresponding to each element in the feature vector. If P points are distributed as neighbours, then the feature vector length is given as $P + 2$, where the additional number corresponds to the variance and mean parameters of the model. The selection of a spherical neighbourhood system enables the extraction of texture from 3D space, allowing for exploration of information contained within the volumetric image. This is dissimilar to most common approaches to texture analysis of 3D images, according to which texture features are extracted from 2D slices of a volumetric image and the texture patterns arranged in multiple directions are ignored. Besides, generating a 3D GMRF model locally at each voxel (rather than from the entire image volume) means that the underlying texture primitives can be captured, resulting in more discriminative features.

The features extracted by the 3D GMRF model capture the local spatial dependencies of a voxel v given its neighbours. These can be employed in various volumetric texture analysis, such as classification, segmentation, and tracking, which will all be introduced in the following chapters.

3.6 Conclusion

This chapter introduces the extension of GMRF to 3D GMRF for the extraction of texture features from volumetric images. The neighbourhood system adopted in this research consists of a spherical neighbourhood with uniformly distributed points. Such a neighbourhood system means that the neighbour indices are maintained under random transformation such as rotations, helping to further improve the discriminative power of the features. Once neighbours are uniformly distributed, the 3D GMRF is used to model the spatial dependences between the central voxel and its neighbours. The 3D GMRF model parameters are then estimated locally at each voxel using MLE, resulting in a feature vector consisting of the estimated parameters. Unlike most 3D texture analysis methods, according to which 3D images are processed based on extracted 2D slices, this method includes a third dimension by extracting features from 3D patches rather than 2D slices. In the following chapters, texture features extracted by 3D GMRF are used to solve a variety of texture problems.

Chapter 4

Volumetric texture classification

4.1 Introduction

In the previous chapter, the extension of a GMRF model to three dimensions is proposed to extract texture features from volumetric images. The 3D GMRF texture features consist of estimated parameters extracted for each voxel and forming the feature vector $\boldsymbol{f}_v = [\boldsymbol{\alpha}_v{}^T, \sigma_v^2, \lambda_v]$.

In this chapter, various descriptors are constructed based on these features and employed in volumetric texture classification. Section 4.2 introduces a set of descriptors built upon the 3D GMRF texture features. The descriptor that is formed by computing the distribution of the features extracted by the 3D GMRF model is presented in Section 4.2.1. Section 4.2.2 introduces the descriptors that are further improved by encoding them into single codes using a binary encoding scheme. In Section 4.3, the descriptors are evaluated using the RFAI synthetic texture database. This evaluation reveals that encoding the extracted features leads to a significant reduction in descriptor size, while preserving the discriminative power of the 3D GMRF features compared with the other methods. In Section 4.4, the proposed descriptors are used as a tool for the diagnosis of COPD. In this section, a brief explanation of this disease is given, and the current methods for COPD diagnosis are presented. Then, a performance evaluation is carried out using a clinical dataset to demonstrate the classification performance of the descriptor proposed herein compared with the other methods. The chapter is concluded in Section 4.5.

4.2 Descriptor construction

In computer vision, image classification refers to the process of associating a desired class label with an image (Nixon et al. 2008). To this end, an image is often represented by a feature vector or descriptor describing its key characteristics, such as its texture. In this research, the local features represented by estimated parameters are extracted for each voxel in the volumetric images, which contain discriminating information about the local texture. The distribution of these features over the volume of the image is then computed to produce a single descriptor describing the volumetric texture of the image. Computing the distribution of local features has proved to be an effective approach for texture classification, as it combines structural and statistical approaches (Ojala, Pietikäinen and Maenpaa 2002; Jianguo Zhang et al. 2007). By modelling the spatial dependencies between the central voxel and its neighbours using 3D GMRF, the underlying primitive texture is detected, and its distribution over the image is estimated by computing a histogram. Altogether, this combination is efficient for texture classification.

In this chapter, four descriptors are proposed based on features extracted by 3D GMRF. The first descriptor is constructed by computing a normalised histogram of each feature in $\boldsymbol{f}_v$ over the entire volume and concatenating them to form a single descriptor. The remaining descriptors are built after decomposing and encoding the estimated $\boldsymbol{\alpha}_v$ parameters of the 3D GMRF into single codes before computing their distribution over the entire volume. The main difference between the descriptors is that the first descriptor treats each α_r (which measures the interaction between a single neighbour and the central voxel) as a single feature, while the others encode the $\boldsymbol{\alpha}_v$ parameter vectors into a single code. These descriptors are explained as follows.

4.2.1 Feature histogram

The first descriptor, denoted as $GMR^{3D}_{P,R}$, is constructed by computing the distribution of each of the estimated parameters α, σ^2, and λ over the entire image volume to produce a single histogram for each parameter. These histograms are then normalised and concatenated to form the descriptor, which is employed as a representation of the entire volume. The complete descriptor is given by:

$$GMRF^{3D}_{P,R} = \{H(\alpha_1)\,, H(\alpha_2)\ \ldots\ , H(\alpha_r)\,, H(\sigma^2)\,, H(\lambda)\} \tag{4.1}$$

Here, $H(.)$ is the normalised histogram of a given parameter over the whole volume, superscript ^{3D} indicates 3D volumetric texture, while P and R refer to the sampling rate of the voxels placed on the sphere used for generating observations and the radius of the sphere, respectively. Figure 4.1 shows an illustration of this descriptor. By varying the values of P and R, the descriptor can capture features at different spatial resolutions.

The size of the $GMR^{3D}_{P,R}$ descriptor depends on the number of voxels sampled on the sphere and the histogram bin settings for each parameter. It is given as follows:

$$|GMRF^{3D}_{P,R}| = (P+2)\times b) \tag{4.2}$$

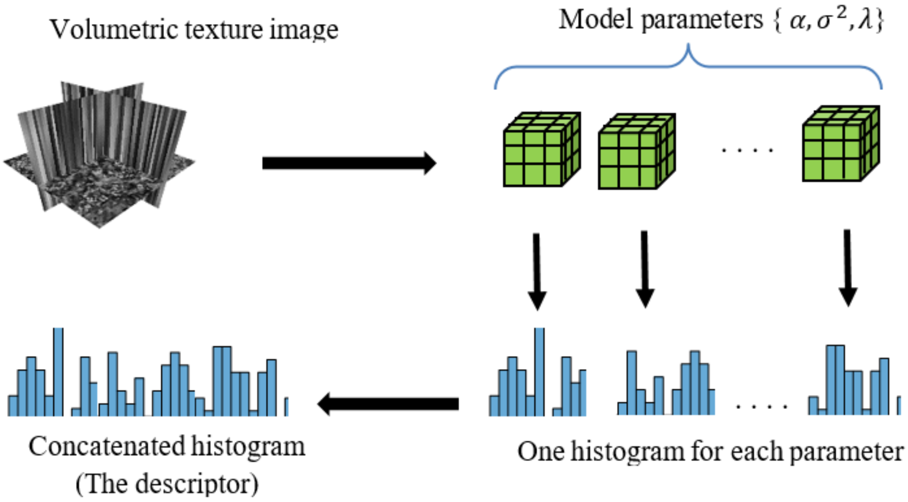


FIGURE 4.1: Construction of $GMR^{3D}_{P,R}$ descriptor

Here, P is the number of voxels, b is the number of histogram bins, and the additional number corresponds to σ^2 and λ. It can be observed that the descriptor size depends entirely on the number of neighbours surrounding the central voxel (i.e., the size of the parameter vector $\boldsymbol{\alpha}_v$).

4.2.2 Binary feature histogram

As mentioned earlier, the feature vector $\boldsymbol{\alpha}_v$ is composed of the estimated parameters α_r that characterize local texture by measuring the strength of the relationship between the central voxel g_v and its neighbour $\boldsymbol{y}_v$. The dimension of $\boldsymbol{\alpha}_v$ is equivalent to the number of neighbours distributed around g_v. Encoding $\boldsymbol{\alpha}_v$ into a single code helps to significantly reduce the computation time required for texture classification without losing its descriptive information. We use the binary encoding scheme proposed by Ojala, Pietikäinen and Harwood (1996) to encode the sign and magnitude of $\boldsymbol{\alpha}_v$. To this end, the vector $\boldsymbol{\alpha}_v$ is first decomposed into two components represented by magnitude $\boldsymbol{m}_v = |\boldsymbol{\alpha}_v|$ and sign $\boldsymbol{s}_v = sign(\boldsymbol{\alpha}_v)$ vectors defined at each voxel v. The next step is to encode $\boldsymbol{m}_v$ and $\boldsymbol{s}_v$ into single codes as follows:

$$\alpha_v^m = \sum_{p=0}^{P-1} t(\boldsymbol{m}_v(p), t_m)2^p \quad t(x, t_m) = \begin{cases} 1, x \geq t_m \\ 0, x < t_m \end{cases} \tag{4.3}$$

$$\alpha_v^s = \sum_{p=0}^{P-1} t(\boldsymbol{s}_v(p))2^p \quad t(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{4.4}$$

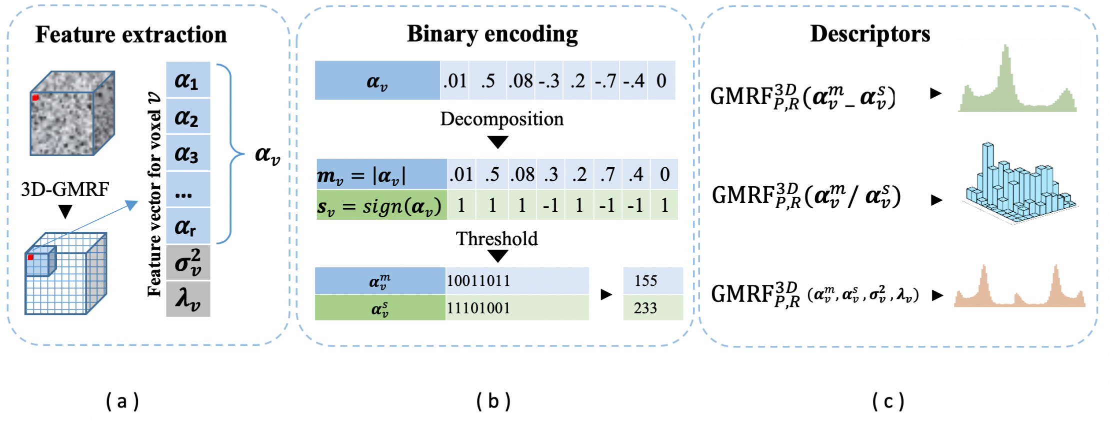


FIGURE 4.2: Overview of the binary encoded descriptors. (a) Features are extracted using 3D-GMRF. (b) $\boldsymbol{\alpha}_v$ is encoded into binary codes. (c) Descriptors are built in various ways.

Here, P is the length of $\boldsymbol{\alpha}_v$, which is equivalent to the total number of neighbours surrounding g_v and t_m is the mean of $\boldsymbol{m}_v$ over the entire volume. The encoded parameters α_v^m and α_v^s can then be combined to construct descriptors by computing either their joint or concatenated normalised histograms. These descriptors are denoted as $\text{GMRF}_{P,R}^{3D}(\alpha_v^m \backslash \alpha_v^s)$ and $\text{GMRF}_{P,R}^{3D}(\alpha_v^m\_\alpha_v^s)$, respectively. An additional descriptor can be constructed by exploiting the other estimated parameters of the 3D GMRF via a concatenation of the normalised histograms of all the extracted features (i.e., $\alpha_v^m, \alpha_v^s, \sigma_v^2, \lambda_v$). This additional descriptor is denoted as $\text{GMRF}_{P,R}^{3D}(\alpha_v^m, \alpha_v^s, \sigma_v^2, \lambda_v)$. Figure 4.2 shows the construction of these descriptors while the details of all descriptors are presented in Table 4.1

4.3 Method evaluation

4.3.1 Dataset, settings, and metrics

All of the proposed descriptors are evaluated on the RFAI synthetic database for volumetric texture (Paulhac, Makris, Ramel et al. 2009), which is described in Section 2.5 of Chapter 2. All of the datasets and their associated subsets except for the rotation subset are used in this evaluation. For each dataset, 50% of each subset is randomly selected for training, while the remainder is kept for testing. This is repeated 100 times and the mean and standard deviation are then computed for all classification accuracies. The neighbourhood scheme is chosen as a sphere with $R = 1$ and $P = 42$ sampled voxels, where the size of the estimation cube Ω_v is selected as $w \times w \times w$ with $w = 4R + 1$, leading to a final size of $5 \times 5 \times 5$. The number of histogram bins is empirically selected to give the best result. The similarities between each of the two samples represented

Table 4.1: Details of the proposed descriptors.

descriptor	construction
$\mathrm{GMRF}^{3D}_{P,R}$	This descriptor is formed by the concatenation of normalised histograms of,the features extracted by 3DGMRF without encoding.
$\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v \backslash \alpha^s_v)$	This descriptor is formed by computing joint normalised histograms of the encoded features α^m_v and α^s_v.
$\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v\_\alpha^s_v)$	This descriptor is formed by computing concatenated normalised histograms of the encoded features α^m_v and α^s_v.
$\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v, \alpha^s_v, \sigma^2_v, \lambda_v)$	This descriptor is formed by the concatenation of normalised histograms of the encoded features in addition to mean and variance histograms.

by the descriptors are estimated using k-nearest-neighbours (kNN; $k = 1$) with L_1 employed as a distance metric. The classification accuracy is calculated as the percentage of correctly classified samples using a leave-one-out strategy to cover all the samples.

4.3.2 Results and discussion

The evaluation results for all the proposed descriptors are presented in Table 4.2. The texture features extracted by the 3D GMRF model successfully classify up to 98.34% of the entire texture database. These results thus demonstrate the discriminative power of the proposed descriptors built by these features. Although each of the four proposed descriptors perform well, $\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v\_\alpha^s_v)$ achieves the highest classification accuracy. One of the advantages of $\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v\_\alpha^s_v)$ is that it is not influenced by increasing the number of neighbours; rather than computing the distribution of every single coefficient, the equivalent parameters of those neighbours (i.e., $\boldsymbol{\alpha}_v$) are encoded into a single code. This means that the dimension of the descriptor depends only on the number of histogram bins, resulting in a stable computation time. This is particularly beneficial when multiresolution analysis is considered, as this kind of analysis requires the sampling of a greater number of neighbours around the central voxel. Figure 4.3 shows the computation time required by each descriptor to classify a dataset with 150 samples. The experiment is implemented using a MATLAB R2019b environment running on an Intel Core i5 3.3 GHz processor with 8 GB of RAM. It can be observed that $\mathrm{GMRF}^{3D}_{P,R}(\alpha^m_v\_\alpha^s_v)$ requires less time than the other descriptors, combining high classification performance with a lower computation time. The $\mathrm{GMRF}^{3D}_{P,R}$ descriptor achieves

Table 4.2: Classification accuracies [%] of the proposed descriptors and the LBP-based method using RFAI datasets.
(a) noise. (b) normal. (c) smooth. (d) subsampling.

Descriptor	Synthetic Texture Dataset																
	Geometric				Fourier				Mixed texture				Interpolated				mean±std
	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d	
$\mathrm{GMRF}^{3D}_{42,1}(\alpha^m_v\_\alpha^s_v)$	99.77	100	99.82	98.92	99.96	99.04	98.68	97.81	99.78	99.92	99.34	100	93.80	95.95	93.78	96.90	**98.34±2.13**
$\mathrm{GMRF}^{3D}_{42,1}(\alpha^m_v\backslash\alpha^s_v)$	99.96	100	98.75	96.97	99.97	99.44	98.69	98.94	100	100	99.58	100	93.97	93.78	90.78	95.95	97.92±2.84
$\mathrm{GMRF}^{3D}_{42,1}\ (\alpha^m_v, \alpha^s_v, \sigma^2_v, \lambda_v)$	98.84	98.88	97.16	98.01	99.04	97.81	94.57	97.30	100	100	99.66	100	93.44	97.90	96.09	94.70	97.71±2.07
$\mathrm{GMRF}^{3D}_{42,1}$	98.80	99.60	98.40	97.60	96.66	97.33	98.66	95.33	97.60	98.80	100	100	96.99	98.66	90.30	97.66	97.65±2.32
LBP (Bhatia et al. 2019)	99.44	99.76	99.68	100	98.40	98.13	97.87	97.78	99.44	99.36	98.72	100	91.21	94.36	89.40	97.79	97.58±3.17

similar classification performance to that of the other descriptors. However, it also suffers from the expensive computational cost required to concatenate the histograms of multiple features, resulting in high-dimensional features. The classification performance of the proposed descriptors is compared with the LBP method based on the decision tree presented by (Bhatia et al. 2019) using the same RFAI dataset. The results, presented in Table 4.2, indicate that the maximum mean classification accuracy achieved by the proposed descriptors is 98.34% ± 2.13 compared with the mean classification accuracy achieved by the LBP-based method (i.e., 97.58% ± 3.17).

While the RFAI dataset contains a variety of texture subsets, the variations within each subset are small, as indicated by the standard deviation of the classification performance across all descriptors. The minor variations within each subset make the dataset insufficiently challenging, highlighting the need to create more comprehensive and complex 3D synthetic datasets.

4.4 Application to chronic obstructive pulmonary disease detection

4.4.1 Chronic obstructive pulmonary disease detection

Chronic obstructive pulmonary disease (COPD) refers to a group of progressive lung diseases and is defined as ‘a common, preventable, and treatable disease characterised by persistent respiratory symptoms and airflow limitation that is due to airway and/or alveolar abnormalities’ (Vogelmeier et al. 2017). There are many signs and symptoms of COPD that can be identified in patients. These include a chronic cough (which is often the first symptom), shortness of breath, wheezing, and sputum production (Vogelmeier et al. 2017). In recent years, COPD has become a serious disease that has gradually emerged on a worldwide scale. Despite important efforts that have been made over the past two decades, some significant issues have yet to be addressed, including the

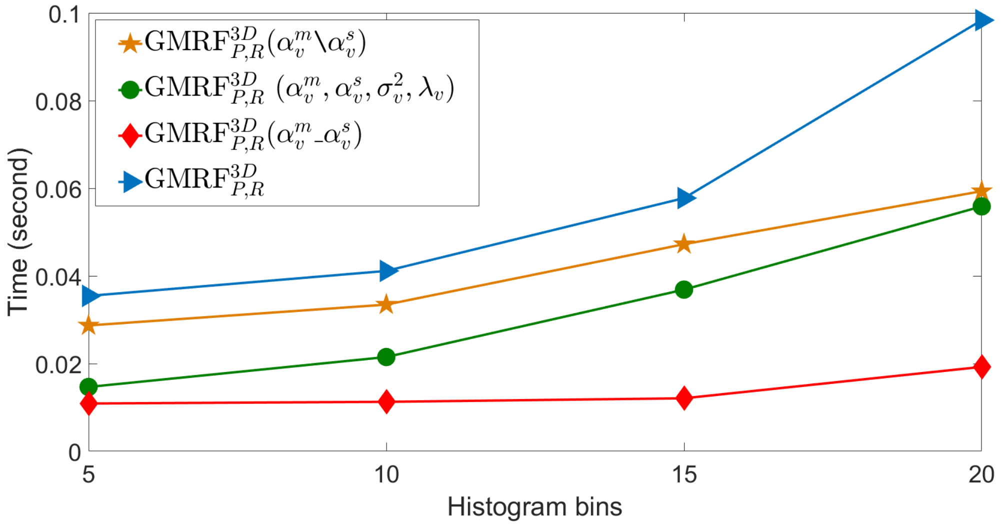


Figure 4.3: Computation time (in seconds) for the proposed descriptors using different numbers of histogram bins.

mechanism of the disease and early diagnosis (Decramer et al. 2012). The principal features of COPD are chronic airflow limitation caused by a parenchymal destruction (i.e., emphysema) and narrow airway disease (obstructive bronchiolitis) (Vogelmeier et al. 2017). Emphysema can be identified by major pathological changes resulting in airspace enlargement and wall destruction in acinus. It is believed that identifying and observing components of the pathological processes could lead to significant clinical benefits (Celli et al. 2004), including early treatment and medical advice. Therefore, developing reliable diagnostic methods to detect the disease in its early stages is a crucial issue.

4.4.2 Diagnostic methods for chronic obstructive pulmonary disease

There are two common approaches to diagnosing COPD: pulmonary function tests (PFTs) and computer-aided diagnosis (CAD).

4.4.2.1 Pulmonary Function Tests

Pulmonary function tests are a group of tests used for lung evaluation. One of these is spirometry, which measures the presence and severity of airflow obstruction (Qaseem et al. 2011). Airflow obstruction not being fully reversible is an indicator of COPD (Qaseem et al. 2011). COPD can be functionally characterised by a loss of lung function measured by forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) (Belchi et al. 2018). While PFT-based measurements are common, inexpensive, and feasible, they also have important limitations, including the fact that they neglect the highly heterogeneous regional pathological changes of COPD (Belchi et al. 2018). Another limitation is that they are not capable of detecting the early stages of COPD

(Sorensen, Shaker et al. 2010; Hackx et al. 2012). Due to these limitations, CAD methods have been developed and used successfully for COPD diagnosis.

4.4.2.2 Computer-aided diagnosis

Computer-aided diagnosis techniques for the quantification of COPD can be broadly categorised into two approaches: intensity-based methods (computed tomography (CT) density) and texture feature methods. Methods based on the intensity histogram of CT attenuation values are popular for the quantification of COPD. Their focus is on deriving measurements of the low attenuation area, a characteristic of COPD, from histograms at a specific threshold. The density mask method is among the early methods that focuses on density histograms as features. Introduced by N. L. Müller et al. (1988), this method aims at measuring attenuation values below a specific threshold, and areas with densities lower than this threshold are considered to be affected by COPD. The suggested threshold values usually fall between –856 and –960 Hounsfield units (HU; Sorensen, Shaker et al. 2010). To determine an optimal threshold, Z. Wang et al. (2013) investigate the optimal threshold for CT assessment of the existence of emphysema in both individual lobes and the entire lung and find that –950 HU is an optimal threshold for density-based emphysema quantification. However, this method suffers from certain disadvantages due to the fact that it considers only independent pixels and discards valuable information, such as the spatial relationships between pixels and structures found on a larger scale (Sorensen, Nielsen et al. 2012). It also depends on a prior selection of the threshold value, making it more sensitive to calibration of and noise in CT images (Sorensen, Nielsen et al. 2012; Gangeh et al. 2010). Another drawback involves overestimating the extent of emphysema due to the presence of other obstructive lung diseases (Park et al. 2008). In addition to this drawback, intensity-based methods become insensitive when only a partial volume or small and dispersed regions are affected.

The limitations of intensity-based methods led to the development of new approaches for COPD quantification. Texture-based approaches are concerned with characterisation of the local structure of the lungs in CT images. Among the earliest texture-based approaches, Uppaluri, Mitsa et al. (1997) introduce a method for evaluating emphysema in CT images that combines statistical and fractal approaches to describe the local structure of regions of interest (ROIs). Later, many more methods are proposed based on texture features (Park et al. 2008; Sorensen, Shaker et al. 2010; Sluimer et al. 2003; Uppaluri, Hoffman, Sonka, Hartley et al. 1999) according to which the descriptor is trained on a set of manually annotated ROIs and used to classify the test set. This exploits the abundance of available information in CT images compared with in intensity-based approaches. Nonetheless, Sorensen, Nielsen et al. (2012) address the limitations associated with the manual labelling of ROIs and propose an entirely data-driven approach for COPD detection in CT images based on texture analysis without the requirement of human intervention. In this method, ROIs are randomly obtained from the lung of a subject and assigned a global label according to PFTs during the acquisition of CT images. Based on the texture features extracted from these ROIs, classification is achieved by measuring the distance between the histograms of the features. One drawback to this approach is that during the patient scan stage ROIs can be partially affected, leading to an increase in noise during ROI classification and meaning that some regions give indications of COPD for a patient that should not be diagnosed with the disease (Cheplygina et al. 2014). One possible solution to this problem is to employ multiple instance learning, according to which a descriptor is constructed for a collection (or "bag") of features and related assumptions are used to describe them. In the case of COPD, for example, an assumption can be made that the subject should be diagnosed with the disease as soon as the lung is found to contain at least one ROI affected by it (Cheplygina et al. 2014).

The importance of texture as a feature has encouraged the development of many texture-based methods, which are regularly employed for COPD diagnosis. These methods have been developed based on various approaches with different feature extraction procedures. In particular, texture features is used by Uppaluri, Mitsa et al. (1997) to quantify emphysema from lung CT images. Further, three statistical methods (the spatial grey-level dependence method (SGLDM), the grey-level run-length method (GLRLM), and the grey-level difference method (GLDM)) are used in a comparative study carried out by Vasconcelos et al. (2010) regarding emphysema diagnosis, in which it is concluded that SGLDM and GLRLM perform better than GLDM. Recently, features extracted by co-occurrence matrices to capture the spatial dependence of grey-level intensities as well as Gaussian derivative methods used to capture structural features have been proposed to automatically detect emphysema without local annotation (Pena et al. 2018).

LBP is used by Sorensen, Shaker et al. (2010) to extract texture features from CT images of lungs, while joint LBP and intensity histograms are computed to describe ROIs. The classification is performed on three categories: normal tissue, centrilobular emphysema, and paraseptal emphysema. Gangeh et al. (2010) use a texton-based approach to classify emphysema in CT images by extracting texton-based features and computing the distributions of those features.

Texture features are extracted by a local distribution of isotropic Gaussian Markov random field (IGMRF) model parameters in which their normalised histograms are constructed as texture descriptors (Dharmagunawardhana et al. 2014). This proposed method, which combines both the statistical and structural properties of the texture, is used for diagnosis and quantification of emphysema, including its subtypes. The classification performance of this method is also improved using the joint features of IGMRF parameter histograms.

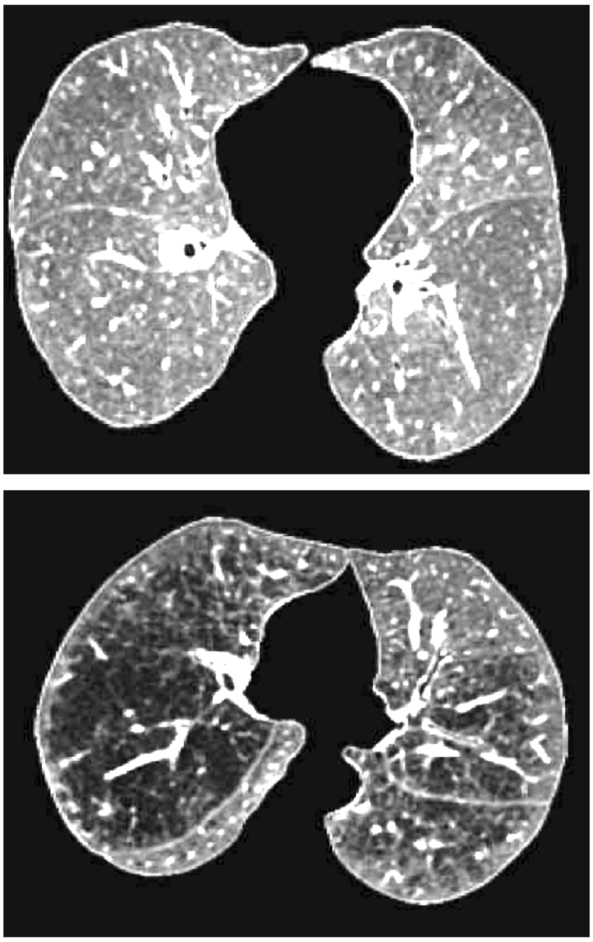


Figure 4.4: Examples of a normal lung (Top) and a lung with COPD (Bottom).

Shamsheyeva et al. (2004) extract texture features from a high-resolution CT (HRCT) image of the lung using wavelet transformation and perform classification with a support vector machine (SVM), combining a discrete wavelet frame (DWF) with grey-level histograms. In addition, Depeursinge, Sage et al. (2007) carry out classification of five selected patterns taken from HRCT images of the lung.

According to the aforementioned methods, texture features are extracted from 2D slices. However, recent advances in medical imaging that offer 3D imaging with richer and more valuable information have created issues related to the feature extraction of volumetric images, as applying 2D methods to 3D images results in a loss of important information (Kovalev, Kruggel et al. 2001). This has led to the extension of many 2D texture methods to 3D methods, such as in the studies by (Park et al. 2008; X. Gao et al. 2010; Citraro et al. 2017), to characterise texture in 3D images, subsequently improving their discriminatory power.

4.4.3 Chronic obstructive pulmonary disease classification using 3D GMRF

As mentioned earlier, the pathological changes caused by COPD affect lung texture appearance in HRCT scans. These textural changes can be used to characterize the disease, meaning that the ability to capture these changes thus leads to the ability to detect COPD from HRCT scans. We select the descriptor that achieves the best classification performance to capture these textural changes. We use a medical dataset of lung HRCT scans consisting of 19 healthy individuals and 13 COPD patients, comprising a total of 32 subjects. Figure 4.4 shows examples of a normal lung and a lung with COPD. The HRCT image volumes obtained from this dataset have a size of $256 \times 256 \times 256$, and our proposed descriptor is employed to distinguish between healthy individuals and COPD patients. The settings and metrics used in this experiment are as previously described. Classification performance is reported based on classification accuracy, sensitivity/recall, and specificity.

A comparative evaluation is carried out here to compare our proposed descriptors with two texture-based methods: a 3D GLCM (Othmen et al. 2013) and a local parameter histogram (LPH; Dharmagunawardhana et al. 2016). In the first method, a set of texture features are derived from a 3D GLCM using different angular directions and distances for each HRCT scan. The second method LPH is a 2D-GMRF-based rotation invariant method and is included in this comparison to investigate the importance of the texture found in 3D space in capturing meaningful information. In this experiment, the LPH variables are set to $p = 8$ and $r = 1$, while the estimation window dimension is set to $w = 5$. This method is applied to each slice in the HRCT scans, and the slice with the highest classification accuracy is selected to ensure a fair comparison.

The results of the classification performance of our proposed method and of the comparative methods are presented in Table 4.3. These results demonstrate that the proposed method achieves the highest classification accuracy (90.62%) in distinguishing healthy individuals from COPD patients compared with the other methods. The comparison of the performance of the proposed method with LPH indicates the significance of texture features extracted from 3D space in providing more discriminative information for distinguishing between samples. Another advantage of our method is that it does not require a large medical dataset for the training stage.

Further investigation is also carried out here to evaluate our method's performance in identifying the differences between different groups in the COPD dataset to understand which groups tend to be at risk for COPD. The differences are identified by measuring the minimum Euclidean distances between the features of pairs of groups. Table 4.4 demonstrates that the healthy smokers group tends to be closer to the COPD group with a minimum Euclidean distance of 0.75 compared with the healthy non-smokers group. It is also noticeable that despite the fact that the healthy smokers are closer

Table 4.3: Comparison of our method with other texture-based methods using the COPD dataset.

Method	Results[%]		
	Accuracy	Sensitivity/Recall	Specificity
$\mathrm{GMRF}^{3D}_{42,1}(\alpha^m_v\_\alpha^s_v)$	**90.6**	**100**	**84.21**
LPH (Dharmagunawardhana et al. 2016)	81.25	100	68.42
3D GLCM (Othmen et al. 2013)	75.00	61.54	82.21

Table 4.4: Differences between COPD, healthy smokers (HS) healthy non-smokers (HNS) groups of the COPD dataset.

Group	COPD-HNS	COPD-HS	HS-HNS
Distance	1.07	0.92	0.75

to the COPD group than the healthy non-smokers, the two groups are still relatively close to each other. This may be because the lungs of healthy smokers have not been significantly affected by COPD and their textural appearance is thus similar to that of the lungs of healthy non-smokers.

4.5 Conclusion

This chapter introduces new descriptors based on the 3D GMRF for volumetric texture classification. Four descriptors are constructed using different techniques based on the distribution of local features extracted by a 3D GMRF model. The first descriptor is formed by computing and concatenating histograms of the 3D GMRF model parameters. The other descriptors are further improved by encoding the 3D GMRF estimated parameters into binary codes and computing their distribution over the entire volumetric image. Using synthetic datasets of volumetric textures, we demonstrate that employing the encoding strategy for 3D GMRF parameters improves classification performance. The features extracted by 3D GMRF are successfully employed for COPD detection by capturing the textural changes used to characterize the disease in HRCT lung scans. Furthermore, we show that the features can indicate healthy smokers' tendency to develop COPD. Unlike most of the existing methods for COPD diagnosis, which extract features from 2D slices of volumetric images, our method involves the extraction of texture features from 3D patches of these images. Our method demonstrates a higher classification performance than the 2D GMRF-based method when the same settings and criteria are used.

Chapter 5

Rotation invariant features for volumetric texture classification

5.1 Introduction

Rotation invariance is fundamental in texture analysis because repeated patterns of texture are arranged at random rotations (Andrearczyk, Fageot et al. 2019). In medical image analysis, the rotation-invariant property of texture is critical in reliably diagnosing diseases based on texture features. For instance, some diseases can be detected by pathological changes that affect the textural appearance of the tissue of interest in the medical image. The local structures or patterns of this texture are arranged not uniquely but in an arbitrary orientation. These randomly rotated patterns require a rotation-invariant approach that can capture these patterns at various rotations.

This chapter proposes a set of rotation-invariant features for volumetric texture image classification based on the 3D GMRF. The mathematical notion of spherical harmonics is employed in the proposed method to produce a set of features that are used to construct a rotation-invariant descriptor. In this chapter, the method is evaluated and compared with other methods in the literature using datasets containing synthetic textures and medical images. The results of our experiments demonstrate the classification performance of the proposed method compared with other state-of-the-art methods. Further, the proposed method is evaluated using a clinical dataset to evaluate its performance in discriminating between healthy individuals and COPD patients. The classification performance of the proposed method is also evaluated in classifying lung nodules in the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI) dataset.

This chapter is organised as follows. Section 5.2 reviews the current methods for 3D image classification and how such methods achieve the rotation invariance. The rotation invariant descriptor is introduced in Section 5.3, while Section 5.4 evaluates the proposed method and discusses the results with comparisons to other methods. The applications of the method to medical image classification are presented in Section 5.5. Section 5.5.1 presents the application of the method to the detection of COPD, while the other application is to lung nodule classification and is presented in Section 5.5.2. Lastly, Section 5.6 concludes the chapter.

5.2 Related work

The main challenge in texture analysis is to deal with the potential changes, such as rotation, scaling, and translation, that occur in texture images. In the natural world, textures can be observed at various scales and orientations; therefore, an optimal texture descriptor should be invariant to transformations to develop efficient methods for texture analysis. Rotation invariance is an essential property of any efficient texture analysis method because many computer vision problems presume that images are captured at random orientations.

Various methods have been developed to achieve rotation invariance for 2D textures. However, extending these methods to 3D textures is far from being a simple task. Optimal rotation invariance requires equidistant sampling, which, though straightforward on the circular neighbourhood of a 2D texture, is complex on a sphere representing a neighbourhood in a 3D texture. There is an additional issue related to the order of the sampled points, which is easily manageable in a circle but difficult to deal with on the surface of a sphere due to its dimensionality. These issues mean that there is a need for the development of new methods for 3D texture analysis with rotation invariance.

Existing methods for 3D texture analysis handle the rotation invariance problem in a variety of ways. A novel method according to which a rotation-invariant descriptor is constructed based on a spherical harmonic representation of 3D shapes is introduced by Kazhdan et al. (2003). This method depends on the amount of energy measured within each frequency of each of a number of concentric spheres of different radii. A function defined in a spherical coordinate system is first decomposed into its harmonics, and these harmonics are then summarised within each frequency. The L_2-norm is then computed for each frequency component to produce 2D descriptors indexed by radius and frequency. Although this method suffers from information loss due to the possibility of having the same harmonic reorientations for two different models, it has been proven to obtain better results in comparison with other descriptors. Despite the fact that this method describes 3D shapes rather than volumetric texture, the idea of using spherical harmonics to achieve rotation invariance is also adopted by Fehr (2007) to construct a rotation-invariant texture descriptor.

In the method proposed by Fehr (2007), spherical harmonics are used for describing 3D regions. The rotation-invariant descriptors in this method are constructed by computing a set of histograms for each voxel on a sphere, which are equivalent to a set of variables that represent the frequency components of LBP codes in a spherical harmonics domain. By measuring the distance between two histograms, the similarity between the two regions can then be estimated. The method proposed by Fehr and Burkhardt (2008) employs spherical harmonics to achieve rotation invariance by computing the minimization of the full correlation over all angles between the grey-values of all points on a spherical surface, given a radius and a volume for the representation. These points are weighted in an arbitrary but fixed order to provide a fast method for correlation across all angles. An extension of the 3D LBP method based on a region-growing algorithm and uniform patterns is proposed by Paulhac, Makris and Ramel (2008). According to this method, a search is undertaken to determine uniform regions located on the surface of a sphere.

Additionally, a method presented by Cid et al. (2017) attempts to classify 3D solid textures based on so-called ‘steerability’, a key property of Riesz wavelets. Steerability indicates that the local response of a randomly rotated image can be obtained analytically from a linear combination of the responses of filter bank components without the requirement of a rotated filter.

Depeursinge, Fageot et al. (2018) investigate the local rotation invariance (LRI) and directional sensitivity (DS) of radiomics features, concluding that optimal radiomics image operators must combine LRI with DS. To achieve this, it is proposed that invariant texture operators based on spherical harmonic wavelets (SHWs) are combined with LRI and DS and then employed for 3D texture classification. The significance of local directional information is also exploited by El khadiri et al. (2018) to develop a local directional ternary pattern (LDTP). According to this method, features are extracted from local ternary pattern (LTP) and combined with the directional features extracted from local directional pattern (LDP) to construct a new descriptor for texture classification.

In addition to these methods, deep convolutional neural networks (CNN) have traditionally achieved rotation invariance by learning through rotation-based data augmentation. However, such methods are costly for 3D images due to the intractable number of rotation arrangements in a 3D space. The rotation invariance has recently been used in CNNs to solve various problems, including texture analysis. Winkels et al. (2019) introduce 3D CNNs with group convolutions (3D G-CNNs), using prior knowledge found in images, such as orientation, to achieve rotation invariance. Worrall et al. (2017) tackle the lack of rotation invariance in CNNs by introducing harmonic networks (h-nets) to achieve rotation invariance. This is carried out by replacing the CNN filter with a circular harmonic that represents all rotated versions of the filter. Andrearczyk and Depeursinge (2018) evaluate the employment of equivariance and invariance to rotations to classify 3D textures, concluding that it is beneficial to classify 3D textures by including built-in equivariance and invariance in CNNs. This does, however, involve considerable computational complexity.

The performance of feature extraction from 3D images using 3D approaches rather than 2D approaches has been investigated in several studies (Paulhac, Makris and Ramel 2008; Yan et al. 2016; Griffiths et al. 2019). In medical image analysis, the development of 3D approaches is crucial because most medical imaging modalities produce volumetric images. However, it is also suggested that 2D slices of 3D volumetric textures be used to provide additional training data for CNNs. This could help to mitigate issues regarding the small amounts of training data available in medical applications (Xie, Jianpeng Zhang and Xia 2019).

Although 3D CNNs are numerically expensive in terms of time and memory, they are able to achieve better results than 2D CNNs because they capture features in 3D space (Yan et al. 2016). In a method presented by Xie, Jianpeng Zhang and Xia (2019), a semi-supervised adversarial classification (SSAC) model is proposed using labelled and unlabelled data. This is employed for benign and malignant lung nodule classification in chest CT scans. According to the method, unlabelled data is used to minimise the need for data annotation by high-skilled professionals, who are not always available. In addition, Hussein et al. (2017) propose 3D CNN multi-task learning for lung nodule characterisation. This method makes use of the volumetric information provided by 3D CT scan images. Such information is lost when 2D slices are used. In their study, the method is then applied for lung nodule classification.

The problem of nodule classification is further investigated by Shen et al. (2017), who introduce a method based on multi-crop CNNs (MC-CNNs) for lung nodule malignancy classification. This method uses a multi-crop pooling strategy that captures salient nodule information and uses it for nodule classification.

Despite the fact that CNNs have made remarkable improvements in image-related tasks, such as classification and segmentation (Xie, Jianpeng Zhang and Xia 2019; Litjens et al. 2017; Xie, Xia et al. 2018), a number of substantial challenges remain that could limit the application of deep-learning-based methods in the medical field. The main issue is that CNNs require a large-scale annotated dataset for training to achieve sufficient performance, and such a dataset rarely exists in the medical domain (Kumar et al. 2016). In medical imaging, it is mostly challenging to obtain a large amount of data due to ethical issues associated with patient privacy (Winkels et al. 2018). The lack of sufficiently annotated data further complicates these challenges, as these datasets need to be accurately annotated by high-skilled experts, which is a time consuming task (Tajbakhsh et al. 2016). In addition, the training stage for CNNs requires extensive computational and memory resources. Although the use of transfer learning resolves issues associated with having a small training dataset, the performance of transfer learning might be influenced by the nature of the specific medical application (Tajbakhsh et al. 2016). Such complications can be avoided by using methods based on handcrafted features, which are beneficial for medical image applications in the absence of extensive training data and powerful hardware.

In this chapter, we propose a rotation-invariant descriptor for volumetric texture classification. Unlike deep-learning-based methods, the proposed method does not require large amounts of training data or powerful hardware, making it suitable for medical image applications, where datasets are usually small. The main aim of this method is to achieve rotation invariance by employing spherical harmonics to remove the effect of rotation on 3D GMRF parameters. The size of the rotation-invariant descriptor proposed here does not increase if more voxels are sampled. This is because the concatenated histograms are computed in terms of spherical harmonic frequencies rather than 3D-GMRF-estimated parameters, leading to a stable computation time. The proposed method is suitable for tasks that deal with texture features of volumetric images having any rotated texture patterns. Since volumetric images are commonly produced and used in medical imaging applications, our method could be used in disease detection to classify discriminative texture features extracted from 3D medical images. In some diseases, pathological changes can be characterised by alterations affecting the texture of tissue, and being able to capture these textural alterations under various rotations could help in detecting the disease. In such cases, the proposed rotation-invariant descriptor can be used to recognise texture patterns that characterise the diseases under any random rotation.

5.3 Rotation-invariant descriptor

If a descriptor does not change under any arbitrary rotation, it is called rotation invariant descriptor. As previously stated, however, constructing such a descriptor is not a trivial task when dealing with a spherical neighbourhood. To achieve rotation invariance, we decompose model parameters as a linear combination of orthogonal spherical harmonics to eliminate the effect of rotation on textures, as rotation tends to alter the interaction parameters α_r. Before applying spherical harmonics, it is important to retain the locations of voxels under any random rotation. This can be achieved by distributing equidistant points on a neighbourhood sphere surrounding the central voxel g_v (see details in Section 3.4.3 of Chapter 3).

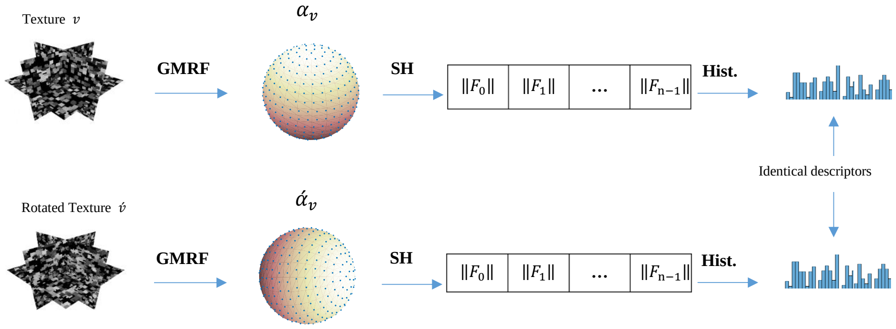


Figure 5.1: Example of constructing rotation invariant descriptors for normal and rotated samples.

5.3.1 Spherical harmonics

Spherical harmonics (SH) are a set of orthogonal functions with which any function in a spherical coordinate system can be linearly decomposed. They are thus analogous to Fourier kernels in Fourier transformation. Transformation with spherical harmonics results in complex-valued coefficients (Green 2003). For a given set of sampled points on a unit sphere, the real $Y^m_{R\ell}(\theta,\phi)$ and imaginary $Y^m_{I\ell}(\theta,\phi)$ parts of the spherical harmonic functions with degree ℓ and order m are defined by:

$$Y^m_{R\ell} = \begin{cases} \sqrt{2}K^m_\ell cos(m\phi)P^m_\ell(cos\theta) & m \neq 0 \\ K^0_\ell P^0_\ell(cos\theta) & m = 0 \end{cases} \tag{5.1}$$

$$Y^m_{I\ell} = \begin{cases} \sqrt{2}K^m_\ell sin(|m\phi|)P^m_\ell(cos\theta) & m \neq 0 \\ K^0_\ell P^0_\ell(cos\theta) & m = 0 \end{cases} \tag{5.2}$$

where $\ell \in \mathbb{R}^+$, $-\ell \leq m \leq \ell$, P^m_ℓ is the associated Legendre polynomial and K^m_ℓ is a scaling factor used to normalise the function and is given by:

$$K^m_\ell = \sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}} \tag{5.3}$$

The projection of functions in spherical coordinates into real and imaginary spherical harmonics are coefficients $C^m_{R\ell}$ and $C^m_{I\ell}$ with degree ℓ and order m. These coefficients are simply computed by integrating the product of the spherical function f and the spherical harmonics function as follows:

$$C^m_{R\ell} = \oint f(\theta,\phi)Y^m_{R\ell}(\theta,\phi)\,d\theta d\phi \tag{5.4}$$

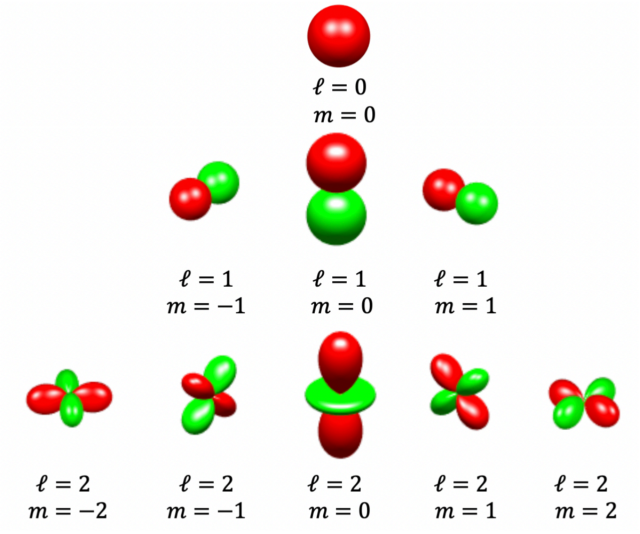


FIGURE 5.2: Illustration of the spherical harmonics of degrees $\ell = \{0, 1, 2\}$ and orders $m = \{-2, -1, 0, 1, 2\}$.

$$C^m_{I\ell} = \oint f(\theta, \phi) Y^m_{I\ell}(\theta, \phi)\, d\theta d\phi \tag{5.5}$$

The discrete approximation of the function f by using the coefficients $C^m_{R\ell}$ and $C^m_{I\ell}$ for a number of bands n is written as follows:

$$\tilde{f}(\theta, \phi) = \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} \Big((C^m_{R\ell} + jC^m_{I\ell})(Y_{R\ell}(\theta, \phi) + jY^m_{I\ell}(\theta, \phi)) \Big) \tag{5.6}$$

where $j = \sqrt{-1}$, ℓ is a positive integer value in the range $[0,n-1]$, and $C^m_{R\ell}$, $C^m_{I\ell}$ are calculated in 5.4 and 5.5, respectively. Figure 5.2 illustrates the spherical harmonics of various degrees ℓ and orders m. Rotation invariance is an important property of spherical harmonics according to which the projection of a function f is identical to any rotation of the projected function $\tilde{f}$ (Kazhdan et al. 2003). Hence, any rotation of the function f does not change its L_2-norm (Fehr 2007). Based on this key property, a rotation-invariant descriptor is proposed that computes the amount of energy contained within each spherical harmonic frequency, which remains constant under any random rotation of the spherical function.

5.3.2 Construction of the 3D GMRF rotation-invariant descriptor

The 3D GMRF model parameters $\boldsymbol{\alpha}_v$ at voxel v measure the influence of voxels y_{v+r}, present on the surface of a spherical neighbourhood, on the central voxel g_v. As previously stated, the rotation of neighbours y_{v+r} results in changes in the parameters α_r. To remove the effect of rotation, a function f is defined at voxel location v to represent the estimated parameters $\boldsymbol{\alpha}_v$ (see Eq. 3.8), which are sampled on the neighbourhood

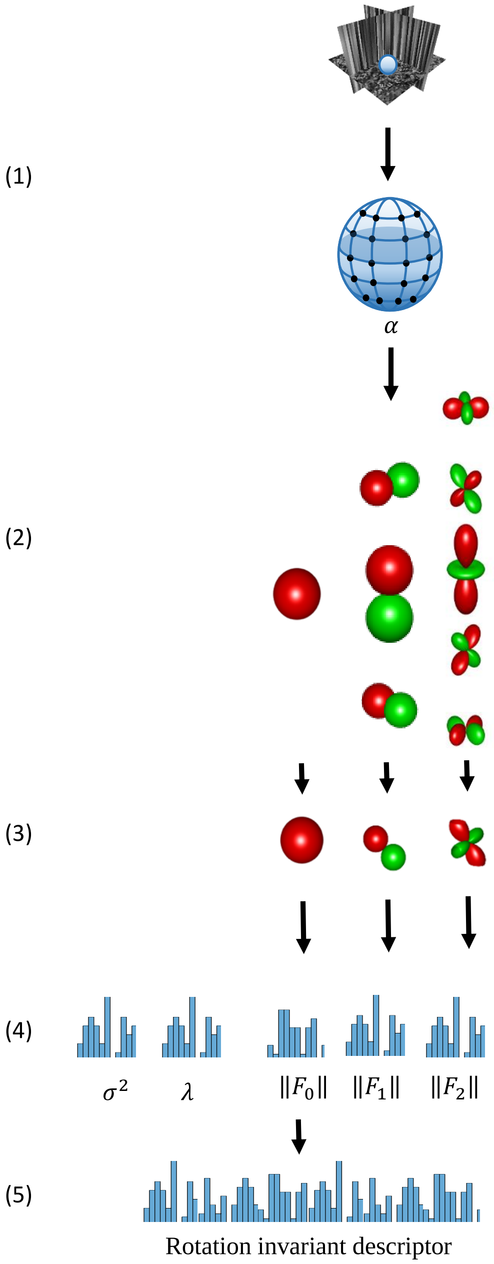


FIGURE 5.3: Construction of the rotation invariant descriptor $GMRF^{3Dri}_{P,R}$. (1) For each voxel in the volume image, discrete points are parametrized on the sphere and parameters α_r are estimated to represent a function in the spherical coordinate system. (2) The function is decomposed into its harmonic frequencies. (3) Within each frequency, the harmonics are accumulated, and a L_2-norm is computed for each frequency component to result in a set of variables for each voxel. (4) A normalised histogram is computed for each variable, including variance σ^2 and mean λ. (5) Histograms are concatenated to construct a rotation invariant descriptor for the whole volume.

sphere. Let $(\theta, \phi) \in \Omega_s$, where Ω_s represents the uniformly distributed points on the surface of a sphere, and let the function f represent the 3D GMRF model parameters $\boldsymbol{\alpha}_v$, the 3D GMRF rotation-invariant descriptor can then be defined by:

$$GMRF_{P,R}^{3Dri} = \{H(\|F_0\|)\,, H(\|F_1\|)\ \ldots\ , H(\|F_{n-1}\|)\,, H(\sigma^2)\,, H(\lambda)\} \tag{5.7}$$

where P is the number of points sampled on the surface of a neighbourhood sphere with radius R. Further, $H(.)$ is the normalised histogram of the descriptor's elements over the entire volume of the sphere, and $\|F_\ell\|$ is the L_2-norm of the frequency component of the function f (Kazhdan et al. 2003), computed by:

$$\|F_\ell\| = \sqrt{\sum_{\theta \in \Omega_s} \sum_{\phi \in \Omega_s} |\, F_\ell(\theta, \phi)\,|^2} \tag{5.8}$$

where $F_\ell(\theta, \phi)$ is calculated as:

$$F_\ell(\theta, \phi) = \sum_{m=-\ell}^{\ell} \left(\left(C_{R\ell}^m + jC_{I\ell}^m \right) \left(Y_{R\ell}(\theta, \phi) + jY_{I\ell}^m(\theta, \phi) \right) \right) \quad , (\theta, \phi) \in \Omega_s \tag{5.9}$$

Each voxel in the volumetric image is composed of a set of elements representing the L_2-norm of the frequency components $\|F_\ell\|$ in addition to the model parameters σ_v^2 and λ_v, as variance and mean are not affected by texture rotation. The reason for eliminating the phase from spherical harmonic coefficients while preserving only their energy is to achieve rotation invariance. For instance, if a texture is rotated randomly, this results in the rotation of its neighbouring $\boldsymbol{y}_v$, which eventually changes its parameter $\boldsymbol{\alpha}_v$. This leads to a parameter $\boldsymbol{\alpha}_v$ that is different from the parameter $\boldsymbol{\alpha}_v$ of the unrotated texture. However, since the rotation of function f that represents the parameter $\boldsymbol{\alpha}_v$ does not change its L_2-norm (i.e., $\|F_\ell\|$), this removes the effect of rotation on the parameter $\boldsymbol{\alpha}_v$ and leads to identical descriptors for a given texture and its rotated version. In practice, this is essential to capture the textures that belong to the same texture class but are arranged at different random rotations. In the example illustrated in Figure 5.1, two volumetric texture patches with different orientations produce identical descriptors. Figure 5.3 shows an illustration of the construction process of the proposed method in greater detail.

5.4 Volumetric texture classification

5.4.1 Metrics, dataset, and parameter selection

The rotation invariant descriptor proposed here is examined using different datasets, metrics, and parameter settings. The RFAI synthetic database (Paulhac, Makris, Ramel

Table 5.1: Classification accuracies [%] of our method and some other methods using RFAI datasets.

Descriptor	Synthetic texture dataset															
	Fourier				Geometric				Interpolated				Mixed texture			
	Rotate	Noise	Smooth	Subsample	Rotate	Noise	Smooth	Subsample	Rotate	Noise	Smooth	Subsample	Rotate	Noise	Smooth	Subsample
A $GMRF^{3Dri}_{42,1}$	**97.3**	100	72.0	**47.0**	**98.0**	**100**	**82.8**	**28.4**	**99.6**	93.0	78.59	49.16	**98.8**	99.6	92.0	26.8
B $GMRF^{3D}_{42,1}$	59.33	–	–	–	70.4	–	–	–	43.81	–	–	–	70.0	–	–	–
C 3D Riesz	97.0	100	**87.0**	45.0	86.0	96.0	42.0	21.0	–	–	–	–	–	–	–	–
D $NI/RD/CI-LBP^{riu3g}$	–	–	–	–	–	–	–	–	95.3	–	–	–	–	–	–	–

Table 5.2: Classification accuracies achieved with various rotation angles using RFAI datasets.

Method	Axis	Classification accuracy [%] for each angle																	
		0°	20°	40°	60°	80°	100°	120°	140°	160°	180°	200°	220°	240°	260○	280°	300°	320°	340°
$GMRF^{3Dri}_{42,1}$	x	100	99.6	96.4	94.4	94.0	89.6	90.4	88.8	87.2	94.4	88.0	86.4	90.0	94.4	95.2	94.8	95.6	99.2
	y	100	99.2	96.4	97.6	96.0	93.6	90.8	88.8	87.6	96.0	90.4	90.0	93.2	96.0	96.8	97.6	98.4	99.2
	z	100	97.6	95.2	94.8	96.0	95.2	92.0	89.2	88.0	92.8	88.4	86.0	87.6	89.2	94.0	95.2	95.2	98.0
$GMRF^{3D}_{42,1}$	x	100	69.6	48.4	42.8	45.6	46.4	40.8	39.2	43.2	67.2	48.0	34.8	41.6	48.0	48.0	47.2	49.6	72.4
	y	100	76.0	59.6	59.6	62.8	56.0	46.8	47.6	51.2	62.8	51.2	48.4	52.8	60.8	59.2	55.6	61.2	76.4
	z	100	65.6	51.2	44.4	44.4	50.8	44.4	38.4	45.6	63.2	40.4	36.8	36.0	43.6	45.2	44.8	47.2	63.2

et al. 2009), described in Section 2.5 of Chapter 2, is used to evaluate the classification performance of the rotation-invariant descriptor. All of the datasets and their associated subsets, including the rotation subset, are used in the evaluation. The similarity between descriptors is estimated using L_1-norm as a distance metric, while the k-nearest neighbours algorithm (kNN) with ($k = 1$) is used to classify the textures. Further, the leave-one-out strategy is used to measure classification accuracy. The optimal number of histogram bins to produce the highest rate of accuracy is also selected; however, the number of bins should not be selected in such a way that it leads to an increase in the descriptor size. Various sample sizes of points P sampled on the surface of the neighbourhood sphere are selected ($P = \{32, 42, 60, 100, 200\}$) to assess the performance of the descriptor, while the radius R of the sphere is assigned to $\{1, 2\}$. The normal (unrotated) subset of each dataset is used as a training set, and each of the other subsets, including those containing randomly rotated textures, is used as a test set.

5.4.2 Results and discussion

In this section, the proposed descriptor $GMRF^{3Dri}_{P,R}$ is evaluated and compared with other classification methods using the RFAI database. The performance of our method with a variety of settings is also evaluated using a dataset selected from the RFAI.

5.4.2.1 Performance evaluation and comparison to other methods

The aim of the first experiment is to ensure that the $GMRF^{3Dri}_{P,R}$ is truly rotation invariant. The method proposed here along with the rotation variant method $GMRF^{3D}_{P,R}$ introduced in Section 4.2.1 of Chapter 4 are both evaluated on each of the rotated subsets of the four datasets. For each dataset, training is performed using the normal subset, and the methods are then tested on the rotated subset. Since the rotation-variant descriptor $GMRF^{3D}_{P,R}$ is constructed by concatenating the histograms of 3D GMRF parameters, any texture rotation leads to changes in parameter order and thus results in unsatisfactory classification performance. The classification accuracy results for both methods, presented in Table 5.1(A-B), demonstrate that the proposed rotation invariant descriptor successfully classifies the rotated set with accuracies ranged from 97.3% to 99.6%. In contrast, the rotation variant method shows reduced performance when the rotated subset is used with accuracies ranging from 43.8% to 70.0%.

Further investigation is carried out by employing additional rotation angles to examine the rotation invariance of the proposed descriptor. The classification performance of $GMRF^{3Dri}_{P,R}$ is evaluated under various rotation angles and is compared with the performance of the rotation-variant method $GMRF^{3D}_{P,R}$. The normal subset of the volumetric texture dataset is first used as a training set, and the two methods are then tested on textures rotated with angles ranging from 0° to 340° around the x, y and z axes.

The results of this experiment, shown in Table 5.2, also indicate that despite the fact that our rotation invariant descriptor is only trained on the normal samples it successfully classifies samples under various rotations. The rotation invariant descriptor proposed here also achieves higher performance than the rotation variant method for all rotation angles except for 0° (at which no rotation is applied). The $GMRF^{3D}_{P,R}$ method has a lower performance in classifying the rotated samples because the descriptors constructed by this method are not rotation-invariant, and classification is highly affected by rotation. The results of these experiments indicate that the $GMRF^{3Dri}_{P,R}$ method proposed here is invariant to rotation.

To further evaluate the classification performance of the $GMRF^{3Dri}_{P,R}$ descriptor, a comparison with two other recently developed texture-based methods is made using the same synthetic dataset. The first method is an LBP-based method proposed by Citraro et al. (2017) called $NI/RD/CI-LBP^{riu3g}$, which combines three descriptors extracted from 3D volumetric textures. The three descriptors describe intensity, radial difference, and contrast, respectively. This method is evaluated using the rotated subset of the interpolated dataset from the RFAI database, consisting of 30 classes, each comprising 10 volumetric images (see Table 2.1 for further description regarding this synthetic dataset). The second method, proposed by Cid et al. (2017), is based on a 3D Riesz wavelet and consists of three approaches for local image orientation estimation. These include one based on multidimensional Riesz components, one based on unidirectional Riesz components, and one based on first-order Riesz components. Our own method is evaluated using all the datasets in the RFAI database and compared with the reported results of the other two methods. All subsets (normal, rotated, noisy, smooth, and subsampling) are considered in this experiment. The normal subset of each dataset is

TABLE 5.3: Classification accuracies [%] for various sampling rates (P), $R = 1$ and two histogram bins values using interpolated dataset of the RFAI database.

P	Classification accuracies [%]	
	Histogram bins = 5	Histogram bins = 10
32	96.99	97.99
42	98.66	99.66
60	98.32	99.66
100	98.32	99.66
200	94.98	93.97

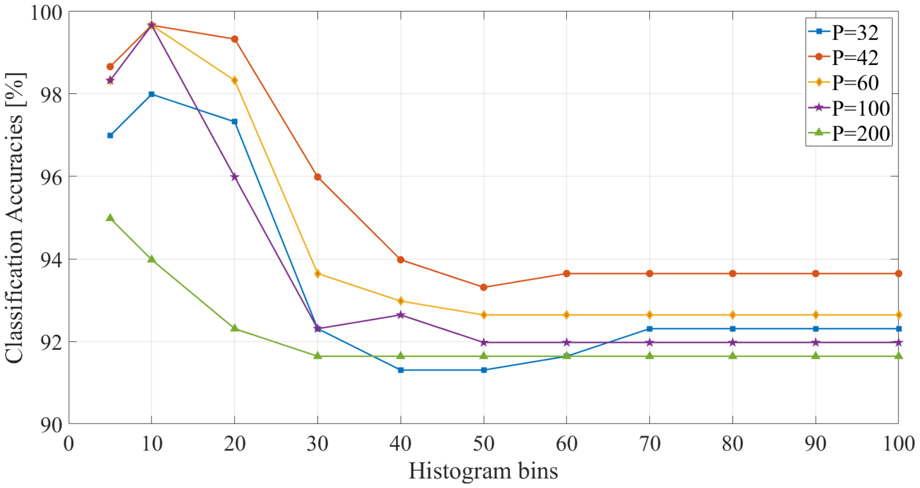


FIGURE 5.4: Classification accuracies [%] achieved for various sampling points (P) and $R = 1$ by employing different numbers of histogram bins.

used as a training set, and the testing is performed on the remaining subsets of each corresponding dataset.

The classification accuracy results presented in Table 5.2(A,C,D) reveal that the proposed $GMRF_{P,R}^{3Dri}$ descriptor achieves a higher classification accuracy (i.e., 99.6%) than $NI/RD/CI - LBP^{riu3g}$ descriptor (i.e., 95.3%) for the interpolated dataset. The proposed descriptor also achieves higher classification accuracies compared with the 3D Riesz wavelet-based method for all datasets except one subset. For the geometric dataset in particular, the proposed descriptor achieves a higher classification accuracy than the 3D Riesz wavelet-based method for all subsets. The descriptor proposed here also performs relatively well for the noisy, smooth, and subsampling subsets data compared with than the 3D Riesz wavelet-based method with an improvement between 3% and 40%. Altogether, our proposed descriptor is favourably competitive for the Fourier dataset and performs better for the geometric dataset compared with the Riesz-wavelet-based method.

The method proposed here is further evaluated using different sampling rates and histogram bin numbers. The classification accuracy results presented in Table 5.3 represent

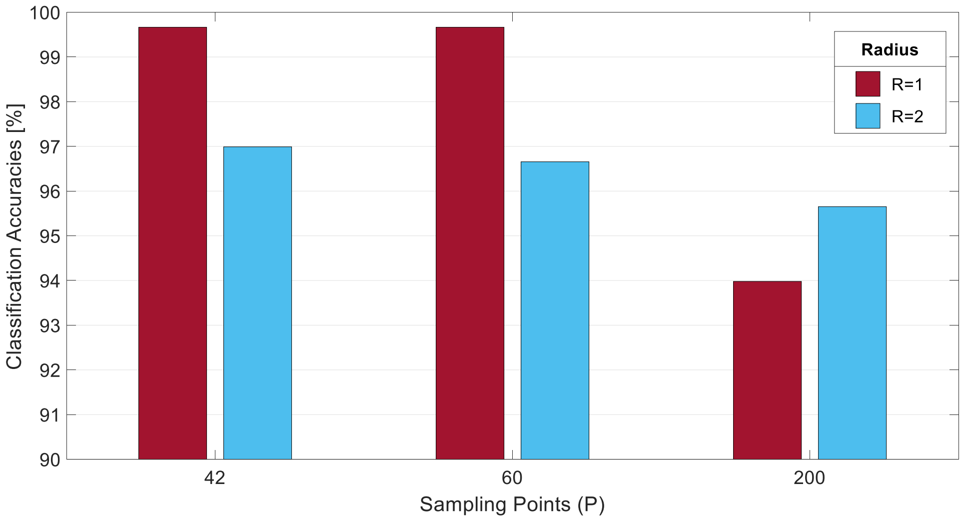


FIGURE 5.5: Classification accuracies [%] of different scales $R = \{1, 2\}$ for different sampling rate $P = \{42, 60, 200\}$ when the number of histogram bins is set to 10.

the classification performance of $GMRF^{3Dri}_{P,R}$ at different sampling rates using the interpolated dataset of the RFAI database. It is found that our descriptor achieves the best results with a sampling rate of $P = \{42, 60, 100\}$ and using 10 histogram bins. With too few or too many sampling points, such as $P = \{32, 200\}$, the performance is slightly decreased. This is because a high sampling rate on a sphere with radius $R = 1$ tends to capture noise, while a low sampling rate leads to the exclusion of fine details, reducing the discriminative power of the descriptor. With only $P = \{42\}$, the descriptor can capture significant details that are sufficient to characterise the local textures with low sensitivity to noise. With $P = \{60\}$, our descriptor also achieves a satisfactory result; however, it requires more computation time compared with $P = \{42\}$.

Figure 5.4 depicts the classification accuracies obtained using different numbers of histogram bins for the interpolated dataset. It is evident that the accuracy is highest when using a histogram with 10 bins and begins to drop gradually with increasing bin number until it reaches 40. From there, the accuracy remains steady at a low value with very slight alterations. It can thus be concluded that the best number of histogram bins is 10 and this leads to a high classification accuracy with a low-dimensional descriptor.

For the proposed method, scale can be controlled by altering the value of the radius R, as higher values of R lead to the ability to capture larger structures. The classification accuracies displayed in Figure 5.5 indicate that $R = \{1\}$ performs better than $R = \{2\}$ on the interpolated dataset except for at a high sampling rate of $P = \{200\}$. As previously explained, this high number of sampling points tends to capture noise when $R = \{1\}$. The relatively low classification accuracy associated with $R = \{2\}$ compared with $R = \{1\}$ is potentially due to the importance of the adjacent neighbours of the central voxel in texture structures. This is because for larger radii, such as $R = \{2\}$, the adjacent neighbours of the central voxel are ignored while the faraway neighbours, which contribute less, are considered. Nevertheless, larger scales can be used to characterise macro-structures or dominant patterns that exist at these scales (Toro et al. 2013).

The size of the descriptor is based on the number of frequency components resulting from the accumulation of spherical harmonics in addition to the variance σ^2 and the mean λ. In other words, given a number of spherical harmonic bands n, the size of the descriptor can be calculated by:

$$|GMRF_{P,R}^{3Dri}| = (n + 2) * b \tag{5.10}$$

Here, b is the number of histogram bins, while the additional terms refer to the variance σ^2 and mean λ. It is obvious that the number of sampling points P does not affect the size of the descriptor. This is because the parameters $\boldsymbol{\alpha}_v$ of the 3D GMRF model are decomposed into a fixed number of spherical harmonic frequency components proportional to the number of bands n. This helps to avoid the issue of having too many sampling points, resulting in a reduction in descriptor size and lowering the computational time required for matching histograms.

The classification performance of the descriptor is reported here through a cumulative match characteristic (CMC) curve. The CMC curve represents the classification accuracies achieved within rank-k where k is the kth nearest neighbour. As can be seen in Figure 5.6, a perfect classification among all datasets can always be achieved within rank 5, or within rank 2 for the Fourier dataset in particular. The classification performance of the descriptor using the mixed texture dataset is measured using the confusion matrix. The confusion matrices displayed in Figure 5.7 indicate that the proposed descriptor has a less number of misclassified samples, especially for rotated, noisy and smooth subsets. However, a higher number of samples are misclassified by the descriptor when the subsampling subset is used. Focusing on the rotation-invariance property, only a few samples within one class are misclassified by the descriptor using the rotate subset, and this is similar to the performance of the descriptor on the noisy subset.

5.5 Application to medical image classification

5.5.1 Application to detection of chronic obstructive pulmonary disease

In these experiments, the proposed descriptor is employed for the detection of COPD. As previously mentioned, this disease is characterised by pathological changes affecting

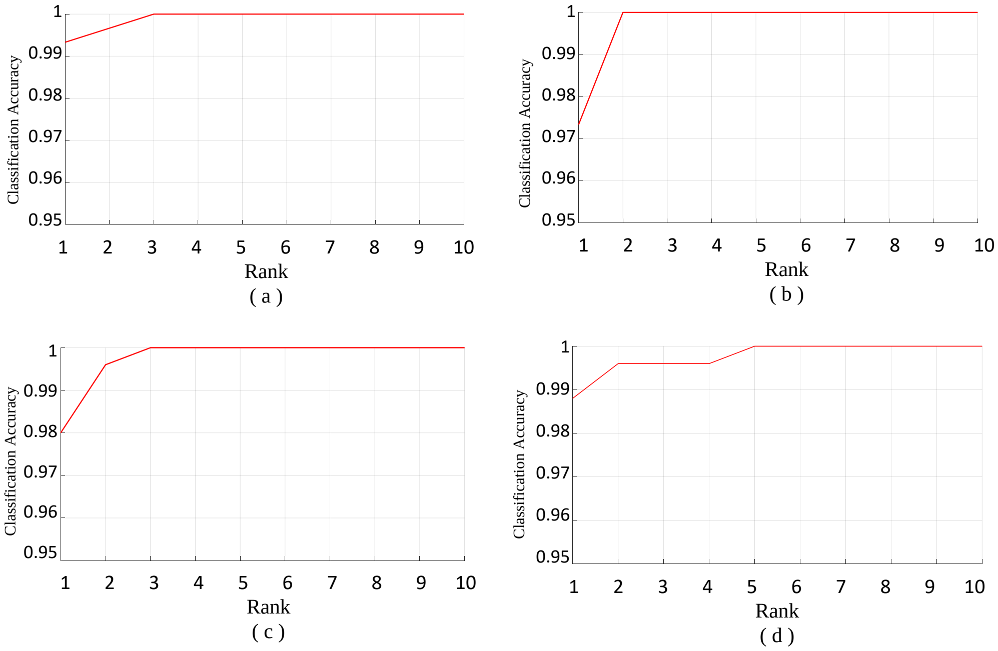


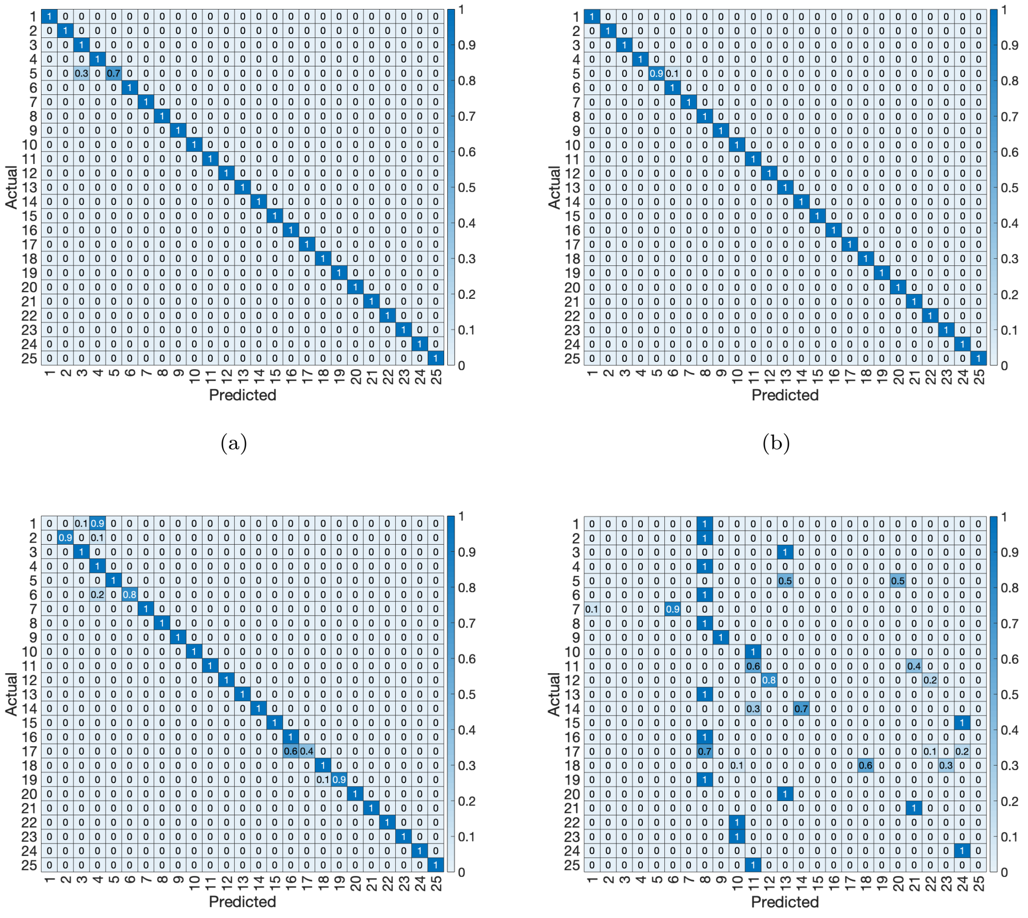
FIGURE 5.6: CMC curves demonstrating the classification performance of our proposed descriptor on different datasets : (a) Interpolated dataset (b) Geometric dataset, (c) Fourier dataset and (d) Mixed texture dataset.

the texture of healthy lungs. The changes in lung texture can be detected by classifying the features extracted from lung CT images to discriminate COPD patients from healthy individuals.

The clinical dataset described in Section 1.4 of Chapter 1 is used in the present experiment. A rotation-invariant descriptor is constructed for each subject by applying our developed method to randomly extracted rotated volumes of interest (VOIs) from HRCT images of lungs (e.g., see 5.8). We conduct three experiments investigating the performance of the proposed method on the COPD dataset. The first experiment, given in the following Section 5.5.1.1, presents a comparison with texture-based methods. In Section 5.5.1.2, our method is compared with intensity-based methods. Finally, in Section 5.5.1.3 a comparison with deep-learning-based methods is made.

5.5.1.1 Comparison with texture-based methods

This experiment aims at comparing our method with texture-feature-based methods using a COPD dataset. Five different texture-based methods are considered along with our previously proposed method $GMRF^{3D}_{P,R}$ to examine the effect of rotation invariance. The five methods include the LPH method based on 2D GMRF (Dharmagunawardhana et al. 2016), an LBP method (Ojala, Pietikäinen and Maenpaa 2002), a 3D GLCM (Haralick et al. 1973; Han et al. 2015), a method based on feature extraction using a bank of Gabor filters (Manjunath et al. 1996), and scale-invariant feature transform



(a)

(b)

(c)

(d)

FIGURE 5.7: Confusion matrices for $GMRF^{3Dri}_{42,1}$ descriptor on the Mixed dataset. (a) Rotate (b) Noise (c) Smooth (d) Subsampling

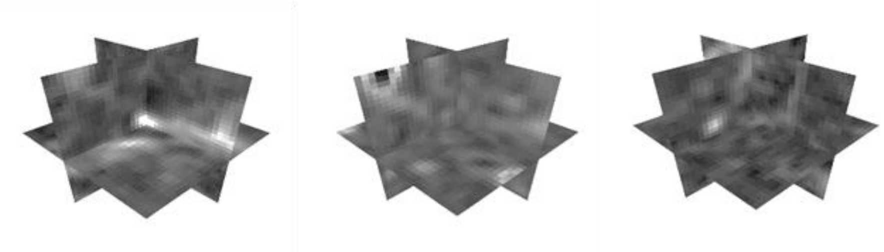


FIGURE 5.8: VOIs sample taken from HRCT images of lungs in the clinical dataset.

(SIFT; Lowe 2004). In the 3D GLCM method, a set of texture features is derived for each VOI from the 3D GLCM using 13 angular directions and four distances. For the LPH and LBP methods, we set $R = 1$ and $P = 8$, while the Gabor filters are computed using six orientations and four scales, as reported by Manjunath et al. (1996). SIFT is a rotation- and scale-invariant method for local feature extraction that has been previously used for various tasks, including texture classification (Y. Yang et al. 2008; Yong Xu et al. 2012). For the SIFT method, a number of key points are detected for each slice in the VOIs, and 128 dimensional descriptors are constructed for these key points. To perform the classification, a normalised global descriptor is obtained by computing the distribution of local descriptors after they are labelled using the k-means clustering algorithm (Y. Yang et al. 2008). The above methods are applied to each slice in the same extracted VOIs, and the best result achieved among all the slices in each of the VOIs is selected to ensure a fair comparison.

The classification accuracies presented in Table 5.4(A–G) demonstrate that the descriptor proposed here achieves the highest classification accuracy (90.63%) in distinguishing between healthy individuals and COPD patients compared with all the other tested methods. Moreover, the outcomes of the comparison between the GMRF-based methods (A and C in Table 5.4), in addition to the performance of the Gabor filters and SIFT support the assumption that the information contained in 3D HRCT images is best described by 3D descriptors. When extracting texture features from 3D images, 3D-based methods are better able to describe 3D textures than 2D-based methods, which process each slice in the volume separately. This separate processing leads to the loss of valuable information that can be accounted for by including a third dimension. Further, since the method shown in Table 5.4(B) is variant to rotation, it is unable to capture different rotated textures and performs worse than the method presented in Table 5.4(A) on the rotated VOIs. Regarding the Gabor-filter-based method shown in Table 5.4(F), besides its low classification performance there is another disadvantage associated with the Gabor filter bank. Namely, it requires filtering the bank of filters with textures rotated under a maximum number of angles to extract features for different possible rotated patterns. Such a filtering process is numerically expressive, time consuming, and is intractable for all possible rotation angles. In contrast, our method can be used to extract rotated features due to its rotation invariance. As can be seen in Table 5.4(A–D), increasing the neighbourhood size (R) is not always beneficial because this incorporates irrelevant voxels that have little influence on the central voxel. Large neighbourhood sizes can also lead to the exclusion of micro-textures that can be detected using a smaller neighbourhood size.

5.5.1.2 Comparison with intensity-based methods

It has already been demonstrated that the texture information extracted from CT images is useful in COPD detection (Park et al. 2008; Uppaluri, Mitsa et al. 1997; Uppaluri,

Table 5.4: Comparison of our method with other handcrafted feature based methods on the COPD dataset for various values of R.

	Method	Classification accuracy [%]		
		R = 1	R = 2	R = 3
A	$GMRF^{3Dri}_{42,1}$	**90.63**	**81.25**	**84.37**
B	$GMRF^{3D}_{42,1}$	81.25	75.00	78.12
C	LPH	75.00	56.25	75.00
D	LBP	78.12	71.88	75.00
E	3D GLCM	75.00	–	–
F	Gabor filters	78.12	–	–
G	SIFT	75.00	–	–
H	Intensity features method	68.75	–	–
I	Density-based method	71.88	–	–

Table 5.5: Comparison of our method with deep learning-based methods on the COPD dataset.

	Method	Classification accuracy [%]
1	$GMRF^{3Dri}_{42,1}$	**90.63**
2	ResNet-50	87.50 ± 0.09
3	AlexNet	85.94 ± .05

Hoffman, Sonka, Hunninghake et al. 1999; Sluimer et al. 2003; Sorensen, Shaker et al. 2010). Additional experiments are conducted here to measure the performance of our method in COPD image analysis when texture features are excluded. The aim of the experiment is to evaluate the discriminative power of texture features for COPD diagnosis through an analysis of 3D lung HRCT scans. The same VOIs previously extracted from the clinical dataset are used in the following experiments.

A. Comparison with features excluding texture information

In this experiment, our proposed method is tested by excluding the texture features $\boldsymbol{\alpha}_v$ and σ_v^2 and using only the intensity features λ_v. This results in the feature vector $\boldsymbol{f}_v = [\lambda_v]$, which leads to the final descriptor $GMRF^{3Dri}_{P,R} = H(\lambda)$. This descriptor is constructed for each of the VOIs, and the same settings and methods described in Section 5.4.1 are followed for $R = \{1\}$. The classification accuracy obtained using only intensity features is 68.75%, as presented in Table 5.4(H). This is much lower than the accuracy achieved with the full-texture model (90.63%), presented in Table 5.4(A), clearly demonstrating the importance of texture features in allowing our descriptor to distinguish between healthy individuals and COPD patients.

B. Comparison with density-based methods

Table 5.6: Comparison of our method with deep learning-based methods on the (LIDC-IDRI) database.

No.	Method	Number of nodules		Results [%]		
		B	M	Accuracy	Sensitivity/Recall	Specificity
A	$GMRF^{3Dri}_{42,1}$	884	514	91.45 ± 0.10	83.48 ± 0.02	**96.07 ± 0.01**
B	$GMRF^{3Dri}_{42,1}$	528	297	**92.63 ± 1.51**	87.27 ± 0.03	95.66 ± 0.01
C	Multi-crop CNN	528	297	87.14	77.00	93.00
D	$GMRF^{3Dri}_{42,1}$	635	509	91.57 ± 1.43	**87.86 ± 0.03**	94.55 ± 0.01
E	3D CNN	635	509	91.26	–	–

In this experiment, our method is compared with a density-based method for COPD detection to emphasise the importance of texture compared with intensity. The presence of COPD (i.e., emphysema) can be detected by identifying low attenuation areas (LAAs) in lung CT images, which are attributable to abnormalities in and destruction of the lung tissues (Muer et al. 1988; Mascalchi et al. 2017). Computing the histogram of LAAs in lung CT images can thus lead to the detection of the disease (Sorensen, Shaker et al. 2010). A variety of measurements can be derived from the histogram of LAAs, including mean density, relative area (RA), and the percentile of LAA distribution (Mascalchi et al. 2017). We select the percentile of LAA distribution as a measurement derived from the histogram of LAAs to compute lung density based only on the intensity values of the VOIs. The percentile of LAA distribution is defined as the distribution of all LAA values in Hounsfield units (HU) under a specific threshold (Mascalchi et al. 2017). The 15th percentile is usually used as an optimal threshold from a practical standpoint (Stoel et al. 2004; Lynch 2014). Thus, it is chosen in this experiment, meaning a threshold of −950 HU is selected for density quantification (Lynch 2014). The classification is performed according to the percentage of voxels under −950 HU in the VOIs. If the 15th percentile of a given VOI is −950 HU or lower, this indicates the presence of emphysema; otherwise, the subject is considered healthy. The classification accuracy of this experiment, as presented in Table 5.4(I), is 71.88%. This result supports the findings of previous research indicating the advantages of texture analysis over intensity analysis for COPD detection.

5.5.1.3 Comparison with deep-learning-based methods

Deep learning approaches, such as those using CNNs, have demonstrated outstanding performance in many fields of machine learning, including image classification. However, the amount of available training data plays a vital role in the performance of CNN models. In this experiment, we aim to compare the performance of our proposed method with CNN-based methods when only a small amount of training data is available. Such a scenario is prevalent in the medical domain, where only a small amount of data may

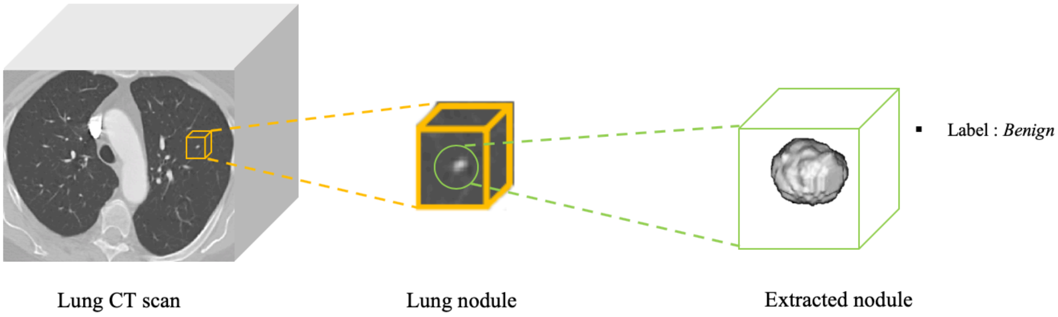


FIGURE 5.9: Extraction process of lung nodules.

be available for one or more of the many reasons discussed earlier. In this experiment, we employ two pretrained CNN models: residual network architecture with 50 layers (ResNet-50) (K. He et al. 2016) and AlexNet (Krizhevsky et al. 2012). These models are trained on the ImageNet database, which has rich feature representations for various image classes. We adapt these networks to our classification problem by replacing the last fully connected layer and the final classification layer with new layers to have a two-class output, matching the number of classes in the COPD dataset. The size of the images is also adjusted so that they fit into the input layer of the networks. 2D slices are then extracted from each of the VOIs across the third dimension and randomly split into training and validation data sets, where 90% of the data is used for training. During training, the images are randomly augmented to increase the size of the training dataset. The adjusted networks are then fine-tuned for 20 epochs with the learning rate set to 0.0001 and employed for COPD classification. The results presented in Table 5.5 are the mean and standard deviation of the classification accuracies of each network running independently 10 times. It can be seen that our method achieves a higher classification accuracy than either the pre-trained ResNet-50 or the pre-trained Alex-Net. However, it is important to mention that the relatively low performance of the deep learning methods is most likely due to the small dataset involved in this experiment. Nevertheless, our proposed method is not affected by the small size of the dataset, indicating that it can be successfully used in medical applications, in which datasets usually have small sizes.

5.5.2 Application to lung nodule classification

The dataset used in this experiment is the Lung Image Database Consortium and Image Database Resource Initiative (LIDC-IDRI), which is publicly accessible for medical imaging research (Armato III et al. 2011). The database contains 1018 clinical chest CT scans, each of which is associated with an XML file that includes detailed information on each case, such as nodule locations and annotations. Each nodule in the LIDC-IDRI database is rated on a 5-point scale with respect to its level of malignancy by up to four experienced thoracic radiologists. We extract the nodules and their associated centres

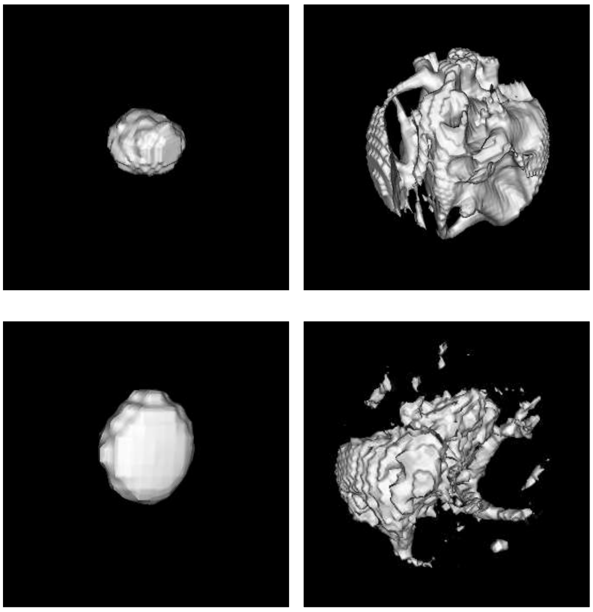


FIGURE 5.10: 3D visualization of benign (left) and malignant (right) lung nodules in chest CT scans.

according to the nodule collection report (A. P. Reeves 2011). The median malignancy level (MML) of each nodule is computed to annotate its level of malignancy. Each nodule with MML > 3 is labelled as malignant, while nodules with MML < 3 are labelled as benign. A nodule with MML $= 3$ is regarded as uncertain and is excluded from this experiment. After removing the nodules with ambiguous MML values as well as those with missing slices, 884 benign nodules and 514 malignant nodules are included in this experiment. A $64 \times 64 \times 64$ cube centred at each extracted nodule is considered to capture the majority of each nodule. We place an adaptive sphere inside the cube to include only the nodule and exclude the surrounding tissue, based on the information provided by (A. P. Reeves 2011). The extraction process of lung nodules is illustrated in Figure 5.9. We then apply our method on extracted patches of the nodules. Figure 5.10 depicts examples of benign (left) and malignant (right) lung nodules in chest CT scans visualised in 3D space. Results are reported as the mean, standard deviation, sensitivity/recall, and specificity of 10-fold cross-validation.

The performance of our proposed method is compared with those of two CNN-based methods (Shen et al. 2017; Hussein et al. 2017). Since their datasets are different with respect to the number of nodules, our dataset is adjusted to match the number of nodules reported for these methods so as to ensure a fair comparison. The reduction of the dataset is performed through a random selection of nodules from the dataset repeated ten times to ensure an unbiased selection of data. Table 5.6 presents the classification performance of the methods involved in this experiment. In row A of this table, it can be seen that our proposed method uses all extracted nodules (884 benign and 514 malignant), while only a part of the dataset is employed for the experiments shown in rows B and D. Table 5.6 shows that the proposed method achieves a maximum accuracy of 92.63% in classifying the lung nodules compared with the Multi-crop CNN and 3D CCN methods that achieve classification accuracies of 87.14% and 91.26% respectively.

5.6 Conclusion

In this chapter, we propose a 3D GMRF-based rotation invariant descriptor for volumetric texture classification that makes use of the rotation invariance of spherical harmonics. According to the method, the function that defines the estimated parameters α_r on the surface of the neighbourhood sphere is decomposed into its spherical harmonics with multiple bands. The harmonics within each band are accumulated, and an L_2-norm is computed for the frequency components. A normalised histogram of each component is computed over the volume, and histograms are also computed for σ^2 and λ. These normalised histograms are then concatenated to construct the rotation-invariant descriptor.

The classification performance of the descriptor is examined by classifying the rotated subsets of various datasets in the RFAI synthetic database for volumetric textures. The experiments show that the proposed descriptor achieves a higher classification accuracy with various sampling rates and a reduced number of histogram bins.

To further demonstrate the classification performance of our descriptor, we present various experiments that compare the performance of our method with that of other state-of-the-art methods on synthetic and medical datasets, including COPD and the LIDC-IDRI medical datasets. The results demonstrate that the classification accuracies achieved by our descriptor are higher in comparison with the other methods. Our descriptor also demonstrates higher performance in comparison with a 2D GMRF-based method in discriminating COPD patients from healthy individuals based on a clinical dataset, indicating the importance of extracting texture features from 3D volumetric data. To further examine the significance of texture features, additional experiments are conducted in which only intensity features are considered. The results of these experiments demonstrate the importance of texture features in COPD detection.

Chapter 6

Volumetric texture segmentation

6.1 Introduction

In this chapter, an adaptive method based on the 3D GMRF is proposed for volumetric texture segmentation. According to the method, a feature vector $\boldsymbol{f}_v$ is extracted for each voxel and used for segmentation. As in Chapter 3, the feature extraction is performed locally using an estimation cube. However, selecting the size of the estimation cube leads to a few fundamental issues related to the uncertainty principle and the inability of the model to capture different texture patterns. These issues are tackled here by employing an adaptive estimation cube with varying size to capture different patterns and minimize the number of voxels that are related to the different texture classes covered by the cube. The feature vectors, which consist of the estimated parameters of the 3D GMRF and form the parameter volume, are employed to segment the volumetric textures. These features are filtered using an adaptive averaging filter. Such an averaging filter improves the segmentation results considerably. The method proposed here is evaluated using a synthetic volumetric texture and compared with other methods for segmentation performance.

In Section 6.2, a variety of existing volumetric segmentation methods are described, while our adaptive volumetric texture segmentation (AVTS) method is introduced in Section 6.3. In Section 6.4, the results of the proposed method are presented and discussed. Section 6.5 presents an evaluation of the proposed method in COPD detection, while Section 6.6 concludes the chapter.

6.2 Related work

Image segmentation is the process of partitioning an image into a set of regions. Texture is one of the various features, along with colour and shape, that can be used for the segmentation. While many methods have been developed to segment textures in 2D images, little attention has been devoted to volumetric texture segmentation (Madabhushi et al. 2003). This is mainly due to the computational challenges involved when including an additional dimension (Aldasoro et al. 2007).

Extracting 2D textures from slices of volumetric images is a common approach to volumetric texture segmentation, but it does not fully make use of the valuable information contained in volumetric data (Depeursinge, Foncubierta-Rodriguez et al. 2014; Madabhushi et al. 2003). Most previous work on volumetric texture analysis has been focused on developing methods for medical image analysis. For example, a method for volumetric texture segmentation based on Fourier domain filtering known as a multiresolution volumetric texture segmentation (M-VTS) is presented by Aldasoro et al. (2007). According to this method, texture measurements are extracted from the Fourier domain of a volumetric image via subband filtering using an orientation pyramid. The Bhattacharyya space is then used to select the most discriminant measurements, producing a feature space. This method uses a multiresolution classification algorithm to classify voxels before boundary refinement is performed. The major disadvantage associated with this method is the involvement of a training stage, which increases its computational cost. Additionally, the method has not been evaluated using images with more than two classes of texture.

An automated segmentation method is proposed by Madabhushi et al. (2003) to segment prostatic adenocarcinoma in MRI data. In this method, the statistical, gradient, and Gabor filter features are computed at multiple scales and orientations to capture the shape, size, and orientation of the tumour. However, this automatic method requires data to be manually labelled by experts, which is not always possible in practical applications.

In a method proposed by Akbari et al. (2012), wavelet-based support vector machines (W-SVMs) are trained to capture kidney textures in 3D MRI images to automatically segment kidneys. These texture features are integrated with geometrical data to differentiate kidneys from their surrounding tissues. Similarly, a kernel support vector machine (KSVM) is used by Zhan et al. (2003) for automatic segmentation of the prostate in 3D ultrasound images. According to this method, texture features are extracted at different scales and orientations based on a statistical shape model using 2D Gabor filter banks applied to orthogonal planes, and the KSVM is used to distinguish the prostate from its surrounding tissues. The employment of SVM-based techniques requires parameter adjustments with high computational costs. Abbasi et al. (2017) use local binary patterns (LBPs) to extract texture features from orthogonal planes of 3D MRI images for the detection of brain tumours.

Luengo, Basham et al. (2016) introduces a method for fast multi-class volume segmentation using supervoxels to represent groups of adjacent voxels. According to the method, regional features are extracted from supervoxels, and an extremely random forests classifier is then trained in a label hierarchy defined by the user to produce a classification output for each layer. This results in final class likelihoods for supervoxels, which are further refined by an MRF model for 3D segmentation. In addition to voxels and supervoxels, the volumetric data in the method proposed by (Luengo, Darrow et al. 2017) includes megavoxels to represent groups of adjacent supervoxels that have similar appearances. This method, called super-region volume segmentation, partitions the volumetric data into hierarchical segmentation layers that are then segmented via user interaction. The major limitation associated with these two methods is that they require user interaction to label the data in every layer, which is not always possible.

Deep learning-based methods have also been proposed for volumetric images segmentation. These include VoxResNet (H. Chen, Dou, Yu et al. 2018), 3D U-Net (Çiçek et al. 2016), and 3D fully convolutional networks (FCNs; H. Chen, Dou, X. Wang et al. 2016). As discussed earlier (Chapter 5), however, the use of deep learning-based methods necessitates a large annotated training dataset. This is generally difficult to obtain in the medical domain because such a dataset requires annotation by high-skilled laborers, which is costly. In addition, the obtainment of medical datasets may be restricted by the ethical approvals required from patients. These restrictions significantly limit the use of deep learning-based segmentation methods in medical applications.

6.3 Adaptive volumetric texture segmentation

6.3.1 Adaptive estimation cube

For volumetric texture classification, the size of the estimation cube Ω_v is empirically selected so that it is constant over the entire texture image. In volumetric texture segmentation, however, a single-size Ω_v sometimes fails to capture relatively large texture patterns. Conversely, increasing the size of Ω_v may create issues with respect to the estimated parameters being over-smoothed. Such a phenomenon occurs when the estimation cube crosses a boundary and covers more than one class of textures, leading to blurred boundaries. The uncertainty principle is another reason for detecting blurred boundaries. Theoretically speaking, to find the exact location of a particular point belonging to a boundary, the estimation cube needs to be as small as possible. In reality, however, a very small estimation cube is not able to capture the texture primitives. To

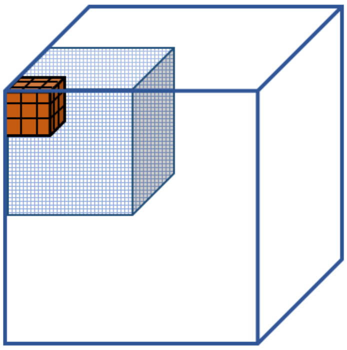


FIGURE 6.1: The neighborhood cube (small) and the estimation cube (medium) inside the entire (large) volumetric image.

capture the texture primitives, the estimation cube needs to be large enough to cover enough 3D texels for a reliable estimation. To tackle these issues, we propose an adaptive estimation cube based on a heuristic rule to select the optimal size of Ω_v before the estimation process is performed. For the heuristic rule, a statistical measure of entropy is employed. It is observed that the entropy of estimation cubes that contain more than one texture are higher than those of estimation cubes containing a single texture. Based on this measurement, a number of estimation cubes (i.e., Ω_v^1 , Ω_v^2 ,..., Ω_v^i) with different sizes are centred at voxel v, and the entropy of each is calculated. The optimal estimation cube is then selected by considering which Ω_v^i has the least entropy as follows:

$$\Omega_v = arg \min_i (H(\Omega_v^i)) \tag{6.1}$$

Here, H is the entropy of each estimation cube, which can be defined as:

$$H(\Omega_v^i) = -\sum_j p_j^i log p_j^i \tag{6.2}$$

where, p is the distribution of the intensity values in the estimation cube Ω_v^i. This procedure deals with the problems associated with boundary blurring by decreasing the size of Ω_v at the boundaries and in this way minimizing the number of voxels covered by Ω_v which are related to different texture classes. It should be noted that this process is not meant to localize boundaries but rather to determine whether a specific Ω_v^i contains more than one texture or not.

6.3.2 Feature extraction

The 3D GMRF model parameters $\boldsymbol{\alpha}_v$, σ_v^2, and λ_v, calculated in Section 3.5 of Chapter 3, are the components of the feature vector $\boldsymbol{f}_v = [\boldsymbol{\alpha}_v{}^T, \sigma_v^2, \lambda_v]$, which contain information

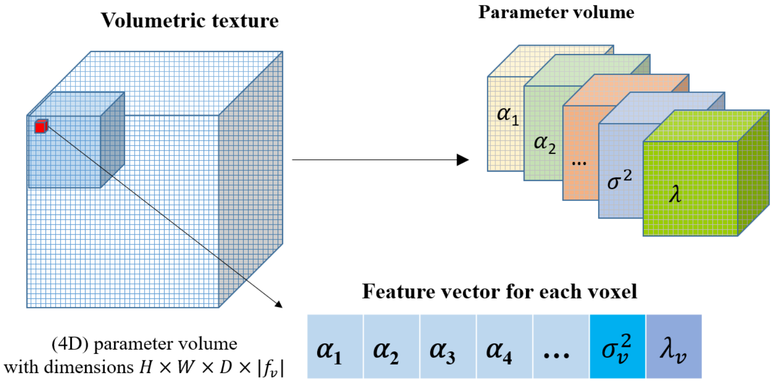


FIGURE 6.2: Graphical illustration of use of the estimation cube to produce a feature vector for each voxel, resulting in parameter volumes.

about local texture. This feature vector is extracted for each voxel at location v using the proposed adaptive estimation cube. Because a cubic neighbourhood system with dimensions $n \times n \times n$ and $n = 3$ is adopted in this chapter, 26 voxels are thus involved in the neighbourhood system which is equivalent in size to the estimated parameter vector $\boldsymbol{\alpha}_v$. Hence, in addition to σ_v^2 and λ_v, the size of the feature vector $\boldsymbol{f}_v$ is $P + 2$, where P is the number of neighbouring voxels. Figure 6.1 depicts the neighbourhood cube that is used to collect neighbouring voxels inside the estimation cube.

6.3.3 Volumetric texture segmentation

In our proposed AVTS method, the normalised texture features extracted by the 3D GMRF are employed for segmentation. First, the 3D GMRF model parameters are calculated for each voxel using the proposed adaptive estimation cube to compose a feature vector $\boldsymbol{f}_v$. Since the feature vector is constructed at each voxel v in a given volume, a four-dimensional (4D) parameter volume is produced with dimensions $H \times W \times D \times |\boldsymbol{f}_v|$, where H, W, and D are the height, width, and depth, respectively (see Figure 6.2). Second, these parameter volumes are filtered by sliding an averaging filter V_m with dimensions $m \times m \times m$ inside the parameter volume to assign each $\boldsymbol{f}_v$ new values. These new parameter values are the feature vectors filtered by V_m. This essential filtering step helps to eliminate isolated voxels that arise either due to the presence of noise or to the nature of the texture structure in a particular region of the image. Finally, k-means clustering is employed to cluster voxels into regions based on the volumetric texture features described by the feature vector $\boldsymbol{f}_v$. Figure 6.3 shows the effect of using the averaging filter V_m to remove isolated voxels during the segmentation process. It can be observed that a greater number of incorrectly segmented voxels are eliminated when a higher V_m is used to filter the feature space. Although a higher V_m

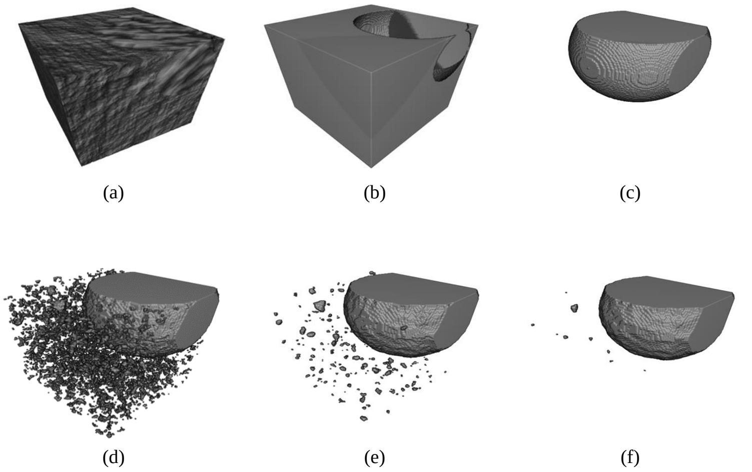


FIGURE 6.3: Illustration of averaging filter effect. (a) volumetric image. (b-c) ground truths. (d) segmented image based on 3D GMRF model parameters without applying an averaging filter. (e-f) segmented images after applying the averaging filter with different sizes to the parameter images.

is beneficial (particularly in the presence of noise), it performs poorly near boundaries. Accordingly, an adaptive approach is proposed to select an optimal value for the averaging filter.

6.3.4 Adaptive averaging filter

The value selected for V_m has a significant impact on the performance of the segmentation. A high V_m helps to remove isolated segmented voxels as well as the effects of noise during the segmentation process, but it also smooths the boundaries between different regions. Overall, manual selection of V_m limits the segmentation performance of the method. One way to address this issue is to employ an adaptive approach according to which the value of V_m varies adaptively according to a specific criterion. Here, we exploited the valuable information encompassed in the 3D GMRF model parameters $\boldsymbol{\alpha}_v$ and σ_v^2 to set up this criterion. In particular, the model parameter σ_v^2 is a good criterion for V_m selection. Given the parameter volume of σ_v^2, the elements collected by V_m will have considerably more variation when V_m contains boundaries, this being due to the inclusion of different texture classes. In addition, the elements will also have lower variation when they belong to a single texture class. Our adaptive averaging filter depending on σ_v^2 is proposed based on these conditions. To achieve this end, multiple averaging filters V_m^i (i.e., $V_m^1, V_m^2, \ldots, V_m^i$) with different sizes are placed at voxel v in a parameter volume of σ_v^2, and the variance is computed for each filter V_m^i. The averaging

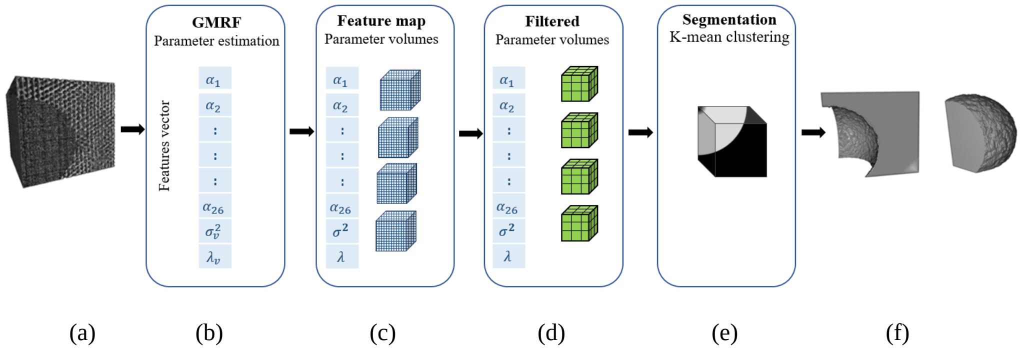


FIGURE 6.4: Overview of AVTS method : (a) Original volumetric image. (b) The 3D GMRF parameters are estimated to construct a feature vector at each voxel which results in parameter volumes. (c) The parameter volumes. (d) The filtered parameter volumes. (e) The segmentation is performed using k-mean clustering algorithm. (f) Results.

filter with the least variance is selected as follows:

$$V_m = arg \min_i (var(V_m^i)) \tag{6.3}$$

This is to ensure that the averaging filter has a maximum value in the homogenous texture regions and a minimum value at the boundaries. Once the appropriate filter is selected, the averaging process is applied before moving to the next location, as described in Section 6.3.3. It is essential to note that the averaging filter is applied to the parameter volume rather than to the given volumetric image. The details of AVTS are illustrated in Figure 6.4. There are a number of advantages to selecting of σ_v^2 over $\boldsymbol{\alpha}_v$ as a criterion for selecting the adaptive averaging filter. In addition to its discriminative power, the parameter σ_v^2 is a scalar for every voxel, making it easy to process at multiple filter sizes.

6.4 Method evaluation

The segmentation performance of our proposed method is evaluated using the RFAI 3D image database (Paulhac, Makris, Ramel et al. 2009). This database contains different categories of volumetric textures for segmentation performance evaluation. Synthetic volumetric texture samples specified for segmentation evaluation and with dimensions $128 \times 128 \times 128$ are selected from the database.

Three categories of images containing two, three, and four texture classes, respectively, are considered in this experiment. Each category contains five texture images, for a total of 15 images. Figure 6.5 exhibits a number of examples of the selected images containing two, three, and four different textures with the corresponding ground truths.

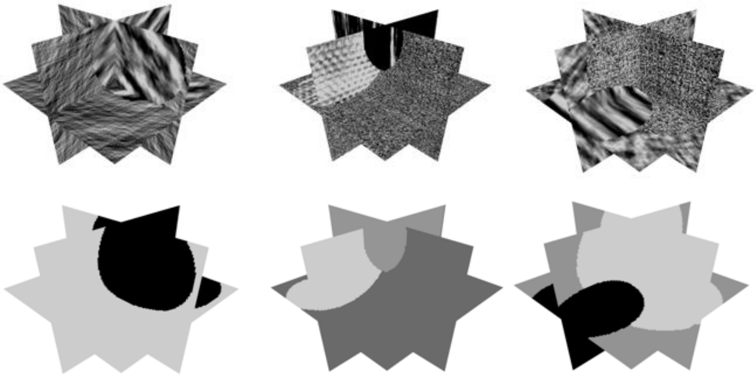


Figure 6.5: Examples of the internal structure of volumetric texture images with corresponding ground truths. Columns from left to right represents images with two, three and four texture classes respectively.

The features $\boldsymbol{f}_v = [{\boldsymbol{\alpha}_v}^T, \sigma_v^2, \lambda_v]$ are extracted at each voxel location using the proposed adaptive estimation cube. The neighbourhood size is set by default to $n = 3$, while the regularization parameter c (see Eq. 3.8) is empirically set to $c = 100$. All of the previous settings are defaults, while the only parameter that requires a manual adjustment is k of the k-means clustering algorithm, which is adjusted to 2, 3, or 4 to cluster the voxels into k separated volumes. The segmentation performance is measured by the error rate E_s and is calculated as the ratio of the number of incorrectly segmented voxels to the total number of voxels in the volume (%). For each texture category, the results are then obtained by computing the mean and standard deviation of the E_s values of all the images.

To evaluate the segmentation performance of our method with respect to the adaptive estimation cube and averaging filter, two experiments are carried out. The first experiment aims to measure the efficacy of the adaptive estimation cube. The 3D GMRF parameters are estimated using a Ω_v with a constant size over the entire sample, and this estimation is compared with that of the AVTS method. We examine three different sizes of Ω_v, $w = 7$, 9, and 11, and an averaging process is employed for both methods. The second experiment aims at examining the effects of the adaptive averaging filter. In this experiment, the 3D GMRF parameters are estimated using the adaptive estimation cube without applying the averaging filter, and this method is compared with the AVTS method.

The segmentation results for our method in the experiments described above are presented in Table 6.1. The results reveal the advantage of employing an adaptive estimation cube during the estimation process. Although the method without the adaptive estimation cube performs well on the images with two texture classes, its performance declines

Table 6.1: Segmentations results in E_s [%]. (1–3) GMRF with adaptive averaging but without the adaptive estimation cube (constant Ω_v size). (4) GMRF with the adaptive estimation cube Ω_v but without adaptive averaging. (5) AVTS (with both the adaptive estimation cube and adaptive averaging).

	Method	Two-class texture	Three-class texture	Four-class texture
1	GMRF, w=7	1.44±0.84	16.33±8.70	38.54±3.61
2	GMRF, w=9	1.693±1.12	6.11±1.95	33.51±4.73
3	GMRF, w=11	2.072±1.45	6.08±1.80	29.83±3.33
4	GMRF, adaptive Ω_v	1.97±1.24	12.18±9.42	37.02±5.77
5	AVTS	1.88±0.98	6.79±2.09	33.47±3.56

when it is generalized to those with three or four texture classes, in which more boundaries exist. Similarly, the averaging process is essential to improve the performance of our segmentation method. An optimal size for the averaging filter must be selected, which can be achieved through the employment of an adaptive averaging filter. The results presented in Table 6.1(4-5) also indicate that the overall performance of our AVTS method is improved by employing both the adaptive estimation cube and the adaptive averaging filter. It should be noted that AVTS has a competitive performance in comparison with the other methods, whose parameter settings are selected manually. Overall, AVTS has the advantage of an adaptive (automatic) scheme for adjusting the size of V_m and Ω_v. Some examples of the segmentation results obtained with AVTS are shown in Figure 6.6

6.4.1 Comparison with other methods

Our method is also compared with the M-VTS method (Aldasoro et al. 2007) in volumetric texture segmentation. AVTS is implemented on the same dataset generated by (Aldasoro et al. 2007) and compared with the previously reported M-VTS results. The results of the comparison, presented in Table 6.2, demonstrate that AVTS achieves segmentation errors that are approximately 10% to 84% lower than the segmentation errors achieved by M-VTS.

Further, our method is compared with a number of other methods for which results using the RFAI database have been previously reported in the literature (Paulhac, Makris, Ramel and Gregoire 2015). These methods include the following: human-understandable features (HUF) (Paulhac, Makris, Ramel and Gregoire 2015), 3D GLCM (Haralick et al. 1973), 3D LBP (Paulhac, Makris and Ramel 2008), and 3D discrete wavelet transform (3D DWT; Jafari-Khouzani et al. 2004). The best results achieved by each method with its optimal parameter settings are presented in Table 6.3 alongside the results obtained using our method. The mean and the standard deviation of the results obtained for all

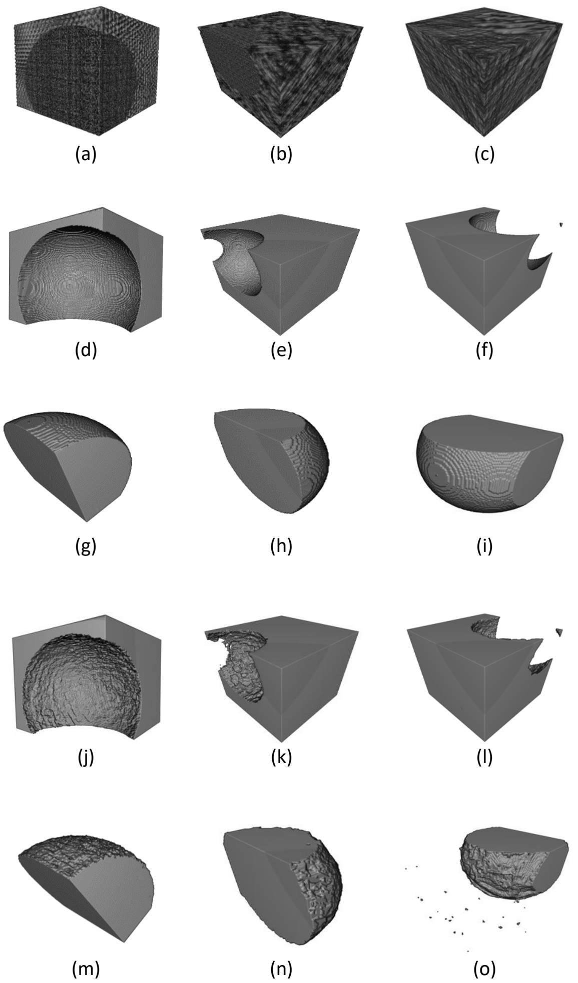


Figure 6.6: Examples of the segmentation results using the AVTS method. (a–c) Original volumetric images. (d–i) Corresponding ground truths. (j–o) Segmentation results.

images in each category are shown. Our method achieves less segmentation error compared with the other methods for 3D images with two or three textures. Using four-class texture images, our method performs just below or close to the other methods. Nevertheless, AVTS is semi-automatic and does not require manual parameter setting (except for the number of textures k) for segmentation. In contrast, the performances of the other methods investigated here depend on the manual adjustment of their parameters, which in some cases produces a high variance in the results.

Table 6.2: Comparison results in E_s [%] of our method against M-VTS.

Datasets	Method	
	M-VTS	AVTS
Gaussian Data	6.2	**0.99**
Oriented Data	3.0	**2.69**

Table 6.3: Comparison of Segmentations performance in E_s [%] of some existing segmentation methods reported in the literature with our proposed method.

Method	Two-class texture	Three-class texture	Four-class texture
$LBP^{riu2}_{26,1}, N = 7$	4.97±9.24	-	-
$LBP^{riu2}_{98,2}, N = 9$	-	7.33±5.42	-
$LBP^{riu2}_{98,2}, N = 7$	-	-	31.18±15.92
$DWT, db2, \beta = 1, N = 5$	4.56±6.08	-	-
$DWT, db2, \beta = 2, N = 5$	-	7.68±4.91	21.882±6.15
$HUF, db2, \beta = 2, N = 5$	3.44±5.06	-	-
$HUF, db2, \beta = 2, N = 5$	-	9.06±3.77	-
$HUF, db2, \beta = 1, N = 9$	-	-	**17.49±12.18**
GLCM	8.56±12.75	14.52±6.02	34.31±11.80
AVTS	**1.88±0.98**	**6.79±2.09**	33.47±3.56

6.5 Application to COPD Detection

In this experiment, we evaluate our method using a clinical dataset of lung images for the detection of COPD. Because COPD can be characterised by changes in lung texture, changes in lung texture can be used to localize regions of the lungs that are affected by COPD. An abnormal subject is selected from the clinical dataset of full-lung high-resolution computed tomography (HRCT) volumetric images with a size of $256\times 256\times 256$ and used in this experiment. Figure 6.7 shows the segmentation results for the selected slices from the axial and coronal viewpoints.

6.6 Conclusion

In this chapter, an adaptive method for volumetric texture segmentation called AVTS is introduced. Based on a 3D GMRF model, estimated parameters are employed as feature vectors to describe local texture. These features are extracted for each voxel in the volume using the proposed adaptive estimation cube. The feature extraction generates parameter volumes that are filtered by an adaptive averaging filter to remove isolated voxels and discriminate between different regions before the segmentation is performed.

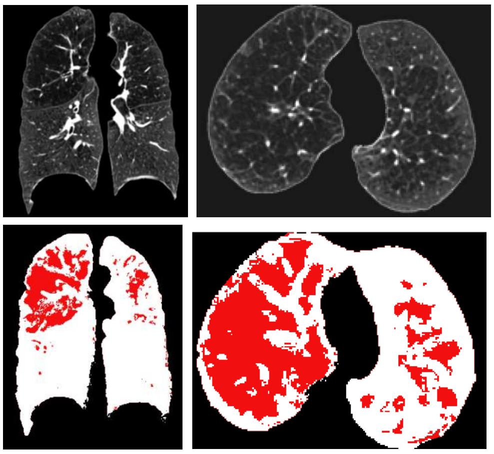


Figure 6.7: Examples of the segmentation results using our method on the clinical dataset. (Top) The original volumetric images of the lung from the (left) axial and (right) coronal viewpoints. (Bottom) Segmentation results. .

The results of the experiments conducted on volumetric textures involving different texture classes demonstrate the performance of AVTS in extracting discriminative features and segmenting volumetric textures. Additional experiments are also carried out to compare AVTS with existing methods from the literature, and the results demonstrate that our method produces competitive results while being a semi-automated method that does not require manual adjustment of the estimation cube and filter sizes.

Chapter 7

Texture-based region tracking

7.1 Introduction

Region tracking becomes a challenging task when dealing with low-contrast textured images, such as cilia video images, in which beating cilia appear as moving texture. Tracking such patterns requires the extraction of features that are capable of discriminating the patterns effectively.

In this chapter, a method for region tracking based on texture features is proposed. The texture features are extracted from a sequence of images using GMRF features. These features are used to track the motion of a given region and extract its trajectory. The proposed method demonstrates the tracking performance depending only on texture features when evaluated on synthetic samples generated for this purpose. The proposed method is also successfully employed to extract the trajectory of ciliary motion, indicating that it could assist in improving analysis of cilia beating behaviour.

The aim of this chapter is to develop a texture-based tracking method to track a textured region and extract its trajectory. Instead of tracking individual pixels, groups of pixels in the underlying texture are tracked based on their texture features, as modelled by GMRF. This method is then used to extract the trajectories of moving cilia, which appear as moving textures in cilia video images.

This chapter is organized as follows. In Section 7.2, diagnostics methods for primary ciliary dyskinesia (PCD) are explained, including automated detection methods. The related work on ciliary motion analysis is also outlined in this section. The proposed method of texture-based region tracking using GMRF is introduced in Section 7.3, while in Section 7.4 , the method is evaluated using a synthetic dataset. The application of the method to ciliary motion analysis is presented in Section 7.5 , and Section 7.6 concludes the chapter.

7.2 Background

7.2.1 Diagnostic methods for primary ciliary dyskinesia

Motile Cilia are a crucial part of the pulmonary defence mechanisms responsible for protecting the respiratory system from harmful bacteria, viruses, and pollutants through a biological system known as mucociliary clearance (Ikegami et al. 2010; Kim et al. 2011). The clearance task is performed by ciliary beating, which leads to the propulsion of mucus. Defects in ciliary beating can lead to the impairment of mucociliary clearance, which is a crucial characteristic of patients with genetically inherited PCD (Morillas et al. 2007; Noone et al. 1999). The diagnosis of PCD can be performed by a number of tests recommended by the European consensus guidelines (Barbato et al. 2009), where the use of a combination of tests to diagnose PCD without a specific gold standard is suggested.

There are two main types of test used to diagnose patients with PCD: screening tests and diagnostic tests (Shapiro et al. 2016). With respect to screening tests, a number of methods are used. For example, a saccharin test is often performed as a screening test for PCD. This test is performed by placing saccharin in the inferior turbinate of the subject, and the time it takes the subject to taste it is used to indicate the presence of PCD (Lucas et al. 2014) . However, this test is difficult to perform in children and produces unreliable results; thus, it is not recommended (Shapiro et al. 2016).

Nasal nitric oxide (nNO) level is another measurement that can be used as a screening test for PCD diagnosis. It is evident that levels of nNo are low in patients diagnosed with PCD. Hence, measuring the level of nNO can lead to the diagnosis of PCD, though it must also be combined with other tests to perform a reliable diagnosis (Shapiro et al. 2016; Lucas et al. 2014).

Diagnostic tests, in contrast, aim at analysing ciliary beat pattern (CBP) and ciliary beat frequency (CBF) to detect PCD. There are a variety of diagnostic test that assess ciliary function and ultrastructure. Transmission electron microscopy (TEM) is used to diagnose PCD through an analysis of ciliary ultrastructure, which can indicate its presence (Papon et al. 2010; Rumman et al. 2017). In this test, cilia are observed at a high magnification ($\times 60,000$) to examine ciliary ultrastructure defects (Papon et al. 2010). According to the method, defects related to microtubular arrangement and dynein arms are counted. These include the absence of the outer or inner (or both) dynein arms, the absence of the central microtubular pair, and the transposition of an outer microtubular doublet (Rumman et al. 2017; Shoemark et al. 2012). The results of this test are determined as the percentage of abnormal cilia among the total number of examined cilia (Papon et al. 2010). However, it has previously been found that PCD can be present without any obvious ultrastructural defects (Stannard et al. 2010; Quinn et al. 2015). Thus, the major limitation associated with this test is that it cannot be used to diagnose patients with PCD in the absence of these defects.

High-speed video microscopy (HSVM) is another diagnostic test that can be used to assess ciliary function based on CBP and CBF. This test is performed by recording live cilia using a high-resolution high-speed digital video camera at a high frame rate (Rumman et al. 2017). CBP and CBF are then analysed to detect abnormal ciliary function by considering the fact that normal cilia beat at 15.4 Hz. However, this test suffers from low sensitivity in detecting abnormal ciliary motion, which can be attributed to its inability to capture the wide frequency distribution found in the biopsies of cilia (Quinn et al. 2015).

Another method that has emerged as a new tool for ciliary analysis is immunofluorescence microscopy (Omran et al. 2009). This method analyses ciliary defects using antibodies that allow for the detection of missing dynein arms along the ciliary axoneme (Omran et al. 2009; Shapiro et al. 2016).

Methods based on computer vision and machine learning have also emerged as tools for automated analysis and detection of potential defects in ciliary motion (Quinn et al. 2015). These could improve current diagnostic tests by allowing experts to analyse ciliary motion more accurately.

7.2.2 Automated detection of PCD

Automated approaches for PCD detection can involve methods of analysing the ciliary motion. Tracking ciliary beating motion (CBM) can be beneficial for evaluating ciliary beating functions and understanding cilia beating behaviour for diagnosis. Evaluating ciliary beating function based on computer vision and image processing techniques includes analysing CBP and CBF. In method proposed by Puybareau et al. (2015), an automated region-based method is proposed to measure ciliary beating frequency using optical flow. According to this method, image sequences of cilia are segmented into consistent regions based on CBF, which is estimated using a dense optical flow.

The method presented by Quinn et al. (2015) considers ciliary motion as a dynamic texture modelled by optical flow. The elemental components are then computed from the optical flow, and digital signatures are constructed from these components for ciliary motion recognition. The method proposed by Kim et al. (2011) uses a combination of image processing and signal processing techniques to estimate CBF from video images of cilia. It is found that using a combination of optical flow and peak detection of CBF signals leads to good results in measuring CBF.

Overall, however, little attention has been dedicated to analysis of CBM based on texture features. Because the patterns of cilia and their surrounding areas are textural, more information can be obtained from ciliary motion using texture features compared with grey scales. Although tracking textured regions within a textured scene is a challenging task, tracing such patterns would assist in providing an understanding of ciliary motion by extracting its trajectory. This in turn could help to diagnose cilia-related diseases, such as PCD. Although optical flow-based methods model the apparent motion of each pixel between two consecutive frames, they do not explicitly track regions (Quinn et al. 2015). Developing a method that is capable of tracking texture regions is essential in understating ciliary motion and behaviour.

7.3 Method

7.3.1 Feature extraction

The texture features are extracted from each 2D image of the image sequences. Each 2D image is defined as a set of grid points on a 2D regualr lattice Ω with sizes $H \times W$ and indexed by $s = (i, j)$, where $\{1 \leq i \leq H, 1 \leq j \leq W\}$. Then, the GMRF model parameters $\boldsymbol{\alpha}_s, \sigma_s^2$ and λ_s are estimated for each pixel s and are used as texture features. These features encompass texture information of the area surrounding the central pixel g_s at location s with $w \times w$ neighborhood. A set of GMRF parameters for each individual pixel s is generated at this stage to form a set of parameter maps referred to as parameter images in this chapter.

The proposed method utilizes the texture features defined by the normalised feature vector $\boldsymbol{f}_s = [{\boldsymbol{\alpha}_s}^T, \sigma_s^2, \lambda_s]$ to track a given region consisting of a set of neighbouring pixels. In a sequence of images, a set of adjacent pixels is assumed to move together with slight variations in brightness, scale, etc. We minimize the difference between the tracked and targeted regions to find the best-matching region and in this way extract its trajectory throughout the frames.

7.3.2 Region tracking

We use a simple block-matching search technique to track the motion of a region based on texture features. The tracking process is applied to the parameter images generated by GMRF rather than the original image. Let I_t and I_{t+1} denote to the parameter images of two consecutive frames t and $t+1$, respectively. Then, let $\Omega_t^p = \{p = (i, j) | p \in I_t\}$ denote the targeted tracking region with dimensions $m \times m$, and let $\Omega_{t+1}^q = \{q = (i, j) | q \in I_{t+1}\}$ denote the search window with dimensions $n \times n$ in the next frame $t + 1$. The initial value of location p in the first frame is specified manually by the user and should be roughly at the centre of the tracked region. The search window Ω_{t+1}^q has $K = n - m + 1$ overlapping blocks index by $k = 1, 2, \ldots, K$. For each current frame t of the parameter

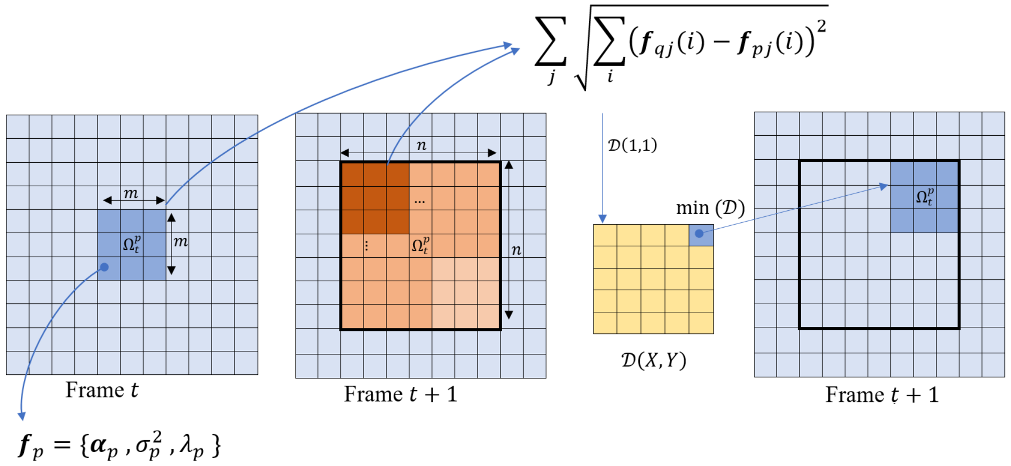


Figure 7.1: Graphical illustration of the region tracking process.

images, we search for the best match for Ω_t^p in the search window Ω_{t+1}^q in the next frame $t+1$ according to the following steps:

1. The similarity between Ω_t^p and the kth block inside Ω_{t+1}^q with the same dimensions as Ω_t^p (i.e., $m \times m$) is estimated by computing the Euclidean distances between the feature vectors $\boldsymbol{f}_p$ and $\boldsymbol{f}_q$ of every two points in Ω_t^p and in the corresponding block inside Ω_{t+1}^q. This estimation is performed as follows:

$$d = \sum_j \sqrt{\sum_i (\boldsymbol{f}_{qj}(i) - \boldsymbol{f}_{pj}(i))^2}, 1 \leq i \leq |\boldsymbol{f}_s|, 1 \leq j \leq m^2 \tag{7.1}$$

2. The distance d is assigned to the distance matrix $D(X,Y)$.
3. The block inside Ω_{t+1}^q is moved by one pixel, and steps 1 and 2 are repeated until the entire search window Ω_{t+1}^q has been covered.
4. Checking all locations in the search window results in the distance matrix D with dimensions $K \times K$. The location of the minimum value in D is the best match for Ω_t^p and thus determines the new location of the tracked region Ω_t^p in the next frame $t+1$.
5. The new location is stored in the trajectory vector T_t.

The previous steps are repeated over all frames, resulting in a trajectory vector of the entire tracked region. This trajectory vector depicts the region motion (i.e., the pattern of the ciliary motion) over all frames. The details of the region tracking process are illustrated in Figure 7.1.

Employing a distance matrix is essential here, as the distance matrix determines the new location of the tracked region in the next frame. It should be noted that the movement

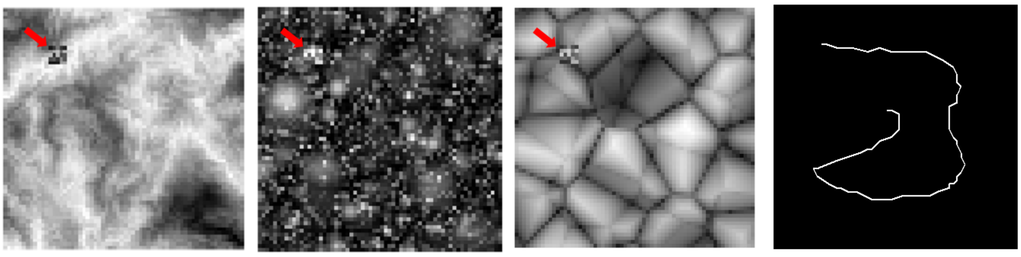


Figure 7.2: Examples of synthetic samples with the ground truth. The red arrow points to the textured region being tracked with size (5 × 5) and placed in different backgrounds.

of the tracked region is not constant over the frames but rather depends on the direction and distance of the minimum value of the distance matrix.

7.4 Method evaluation

Our proposed method is first evaluated using synthetic samples. The synthetic samples are generated from the interpolated dataset of the RFAI texture database (Paulhac, Makris, Ramel et al. 2009), and the size of each sample is $64 \times 64 \times 64$. For each selected sample from the dataset, a textured region with size 5×5 is placed in different positions in the third dimension through all 64 images, producing a moving textured region with different textured backgrounds. The positions of the textured region are recorded as the ground truth of the trajectory of the region. Figure 7.2 shows examples of samples generated with the ground-truth trajectory of the region extracted from all the frames. It is important to notice that the 64 images that form the 3D texture are not identical but differ slightly from location to location in the third dimension, meaning that they can be considered to have dynamic texture.

7.4.1 Performance measurement

The tracking performance of our method is measured by the dissimilarity in pixels between the ground-truth trajectory T_g and the tracked region trajectory T_t. This is determined by computing the distance between T_t and T_g through all frames t as follows:

$$d_t = \sqrt{(p_t - x_t)^2 + (q_t - y_t)^2} \tag{7.2}$$

Where (p_t, q_t) and (x_t, y_t) are the positions of T_t and T_g (Needham et al. 2003). Next, both the mean and standard deviation of d_t are considered to evaluate the dissimilarity between the trajectories, where a mean distance close to zero indicates perfect similarity. In each experiment carried out here, the average is computed for all samples.

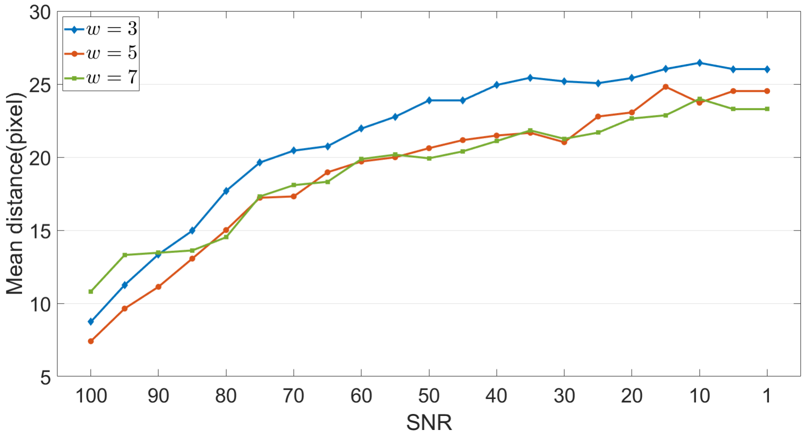


FIGURE 7.3: Tracking results for different SNR values using $w = \{3, 5, 7\}$.

TABLE 7.1: Tracking results (mean and standard deviation) for the synthetic samples using different neighbourhood window sizes.

Neighborhood	$w = 3$	$w = 5$	$w = 7$
Distance	**2.53± 1.54**	8.20 ± 3.37	19.71 ± 7.70

7.4.2 Experiments

Our proposed method is evaluated in both synthetic and medical samples using different neighbourhood window sizes ($w \times w$) during the parameter estimation stage ($w = 3$, 5, and 7). The dimensions of the tracked region are 5 × 5, while the search window dimensions are set as 21 × 21. When the method is applied to the cilia medical sample, these dimensions are changed to 29 × 29 and 31 × 31 for the tracked region and search window, respectively, to better capture the larger-scale structure of the cilia beating patterns.

7.4.3 Results and discussion

Table 7.1 presents the results of our method on synthetic samples using different w values. Regarding the synthetic samples, although the tracked region and the scene both have texture patterns, our method has the best target-region tracking performance when $w = 3$. During the parameter estimation, the value of w determines the scale and patterns of the underlying textures that are captured; selecting an appropriate value for w thus depends on the nature of texture. Nevertheless, altering the size of the tracked region during region tracking affects the results. For instance, when $w = 5$ is chosen enlarging the tracked region from 5 × 5 to 9 × 9 leads to better results (i.e., the mean distance decreases from 8.20 ± 3.37 to 5.5 ± 2.23). In the case of $w = 7$, however, enlarging the tracked region from 5×5 to 31×31 leads to a reduced mean distance from 19.71 ± 7.70 to 9.37 ± 3.68.

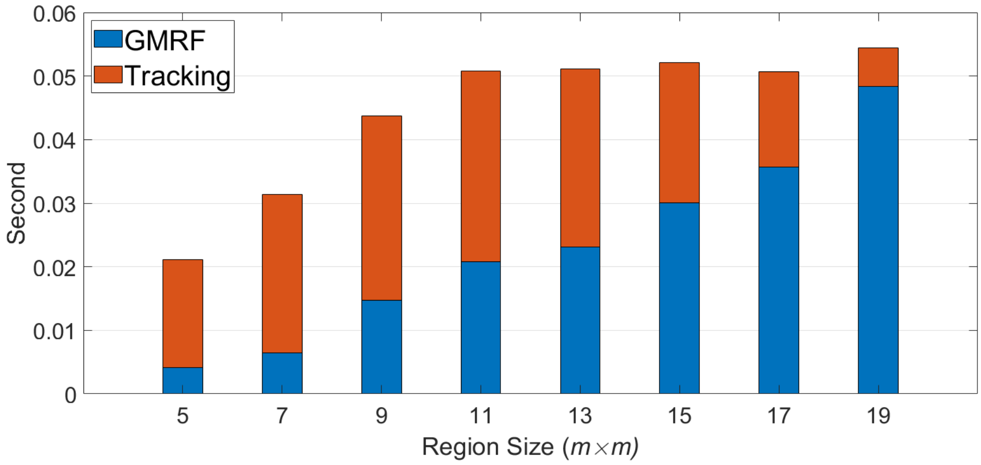


FIGURE 7.4: Computation time (in seconds) of our method for regions of different sizes and $w = 3$.

Further, the method is examined in the presence of noise. Different levels of Gaussian noise corresponding to signal-to-noise ratios (SNR), i.e., 100, 95, ..., 1., are added to the synthetic samples. The SNR is calculated as the ratio of the image power to the noise power. The results for different w values are plotted in Figure 7.3.

Computation time is examined to analyse the efficiency of our proposed method. The method is implemented using the MATLAB R2018a environment running on a 3.3 GHz Intel Core i5 processor with 8 GB of RAM. Figure 7.4 shows the computation times for feature extraction and region tracking for tracked regions with different sizes and with a constant search window size of 21×21 and $w = 3$. It can be observed that most of the computation time is consumed during the parameter estimation stage. This is because the GMRF model is generated for each pixel to increase the discriminatory power of the model, while the computation time during the region tracking is relatively low and not highly influenced by the size of the tracked region.

7.5 Application to ciliary motion analysis

The method proposed here is used to track ciliary motion and extract its trajectory. A video image of beating cilia with size $512 \times 512 \times 512$ frames is used as a medical sample. This video is pre-processed to extract the regions of interest representing the areas where cilia are beating.

7.5.1 Pre-processing of the cilia video image

Our goal is to use the video sample to track the ciliary motion and extract its trajectory. These cilia are found in some regions in the sample and need to be segmented and

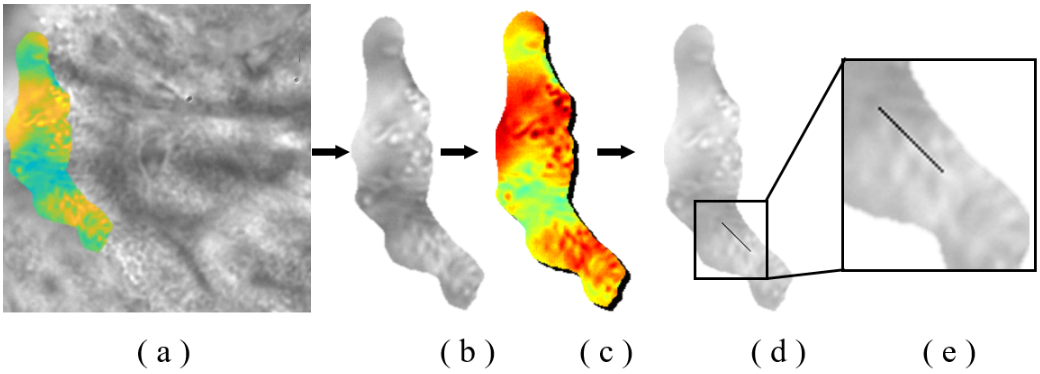


FIGURE 7.5: (a) a selected frame from the video with area of beating cilia isolated by optical flow,(b) the region of interest where cilia motion to tracked, (c) a parameter images resulted from GMRF model, (d) cilia region tracking, (e) extracting the cilia motion trajectory.

TABLE 7.2: Tracking results (mean and standard deviation) using different sizes of neighborhood window on the medical sample.

Neighborhood	$w = 3$	$w = 5$	$w = 7$
Distance	21.48 ± 8.45	**2.80± 5.59**	4.72 ± 4.0

extracted before their motion is tracked. However, camera calibration is a major issue with respect to the video of ciliary motion as it adds additional movement to the video. It is observed that the areas with ciliary motion have high magnitudes of motion compared with the other areas. Therefore, this observation can be used to segment areas with ciliary motion by using the optical flow of the image sequences. The optical flow is estimated for every two consecutive frames using the Horn–Schunck method (Horn et al. 1981), and the magnitude components of the optical flow over all frames are accumulated. By thresholding the accumulated magnitude, only the regions with large magnitude (i.e., the beating cilia areas) are then retained, while the unwanted regions are removed. This step produces only regions with beating cilia. It also helps to reduce the computation time required for the GMRF parameter estimation, as the GMRF models are then generated only for the regions of interest rather than the whole sample.

7.5.2 Ciliary motion tracking

In this experiment, the proposed method is applied to the cilia video clip as a medical sample. The results of this experiment using different neighbourhood window sizes are presented in Table 7.2. It can be seen that the best result is obtained by employing a neighbourhood window size of $w = 5$, although a reasonable performance is also obtained with $w = 7$. When $w = 3$, however, the tracking performance deteriorates, which could be due to the inability of the GMRF model to capture the structure of the underlying texture of the cilia using a small neighbourhood window. However, a larger neighbourhood window tends to lead to the inclusion of irrelevant pixels during parameter estimation, reducing the discriminatory power of the model and decreasing the tracking performance. Figure 7.5 depicts the steps used in extracting the ciliary motion trajectory. The ciliary motion in this sample appears as a straight-line, representing its trajectory.

7.6 Conclusion

In this chapter, a new method for region tracking based on texture features is proposed for region tracking and the trajectory extraction of ciliary motion. The experiments carried out in this chapter demonstrate the performance of the GMRF model in extracting texture features from images for region tracking. These features are implemented via a block-matching search technique, which is employed here to track the region movements and extract the trajectory of the motion. Our proposed method is first examined on synthetic samples generated for method evaluation and then on a medical sample representing cilia. The outcomes of these experiments demonstrate the capability of texture features extracted by GMRF to track a given region, especially when challenging issues exist, such as tracking a textured region in a textured scene.

Chapter 8

Conclusion and future work

8.1 Conclusion

In this research, a new method for extracting texture features from volumetric images is developed based on extended 3D GMRFs. The extracted texture features are then employed for volumetric texture classification, segmentation and texture region tracking. This method exploits the rich information contained in volumetric data by employing a 3D neighbourhood system to explore 3D texture rather than using a slice-by-slice approach.

Chapter 2 discusses the significance of volumetric texture for representing the valuable information contained in 3D images, including medical images. A number of current approaches to extracting volumetric texture features from 3D images are presented, and it is found that the most common approach to volumetric texture analysis is the slice-by-slice method, which cannot be used to explore the rich details provided by volumetric data. This chapter also outlines a number of different neighbourhood configurations that can be used to analyse volumetric texture. It is concluded that 3D neighbourhoods are more capable of extracting texture from volumetric images because they are able to represent various 3D texture primitives.

The extension of GMRFs to 3D GMRFs for the purpose of extracting texture features from volumetric images is introduced in Chapter 3. This method employs locally estimated 3DGMRF model parameters as texture features using a 3D neighbourhood system. The local estimation is performed using an estimation cube to ensure that local texture features are captured.

In Chapter 4, these texture features are used to construct a variety of descriptors that are then used to represent an entire volumetric image. These descriptors are built by computing the distribution of local features over the whole volumetric image, combining detection of the underlying primitive texture with that of its distribution throughout the 3D image to provide better discrimination between samples. Further, a binary encoding scheme is used to encode a feature vector that represents the local features in a single code. This approach helps to reduce the computation time required for matching histograms while preserving the discriminative power of the descriptor. The novel descriptors proposed in this chapter demonstrate higher classification accuracy compared with other methods when evaluated on a synthetic dataset. The proposed method is also used as a diagnostic tool for distinguishing between healthy individuals and those affected by COPD. It is also used to measure the differences between different groups in a COPD dataset to understand which group tends to be at risk for COPD. The investigation reveals that the healthy smokers group tends to overlap more with the COPD group compared with the healthy non-smokers group.

The method presented in Chapter 4 is further improved to capture the textural changes that characterise COPD, and it demonstrates superior discrimination between healthy and COPD-affected subjects by detecting texture patterns under various rotations. This improvement is presented in Chapter 5 via the introduction of a rotation invariant descriptor based on spherical harmonics. The descriptor demonstrates improved classification performance compared with other state-of-the-art 3D texture descriptors for both a synthetic dataset and two medical datasets. The higher classification performance achieved by the proposed method on 3D images compared with the 2D method reveals that volumetric texture is best described by 3D descriptors due to their capability to explore 3D texture primitives.

In Chapter 6, the texture features extracted by 3D GMRF are used for volumetric texture segmentation. This chapter discusses issues related to the necessity of adjusting variables during parameter estimation and average filtering. These issues are solved using an adaptive approach that automatically selects the optimal settings for the estimation cube and averaging filter. This method is evaluated on a synthetic dataset for volumetric texture, revealing superb segmentation results compared with other methods. It is also applied to a lung computed tomography (CT) scan to detect the parts of the lung affected by COPD.

Chapter 7 discusses a number of challenges to textured region tracking, such as the ciliary motion that appears as moving texture in video images of cilia. In particular, the texture features extracted by the 3D GMRF are used to track the motion of a given region and extract its trajectory. The evaluation of the proposed method on a synthetic dataset demonstrates the capability of the extracted texture features to track a textured region and obtain its trajectory. The proposed tracking method is then used to track the ciliary motion of motile cilia and extract their trajectories through a video image of cilia. The outcome of this method could provide a way to improve the analysis of cilia beating behaviour, which is a crucial identifier of PCD.

The applications suited to our proposed method generally involve classification and segmentation tasks for volumetric data with texture characteristics. Medical image analysis is thus one of the possible applications for the methods proposed in this research. The fact that our proposed method does not require a large annotated dataset to achieve promising results makes it further suitable for medical applications, in which large datasets are not often available. The key features of the proposed method, including its rotation invariance, reduced dimensionally, 3D neighborhood configuration, and texture-based region tracking, indicate that it could be used in clinical applications as a diagnostic tool for a variety of diseases that are characterised by textural changes.

8.2 Future work

Not only does our research introduce novel methods that contribute to improving volumetric texture analysis, but it also raises issues that suggest a number of avenues for future research. Some of these possible research directions are as follows:

8.2.1 Inhomogeneous volumetric texture analysis

Detecting texture inhomogeneity in volumetric texture is a challenging task, especially considering its computational complexity. Improving 3D GMRF models to detect texture inhomogeneity would be highly beneficial for the accurate detection of a number of diseases, including COPD. Because the regional pathological changes of COPD are highly heterogeneous (Belchi et al. 2018), this improvement could reliably assist in detecting this disease. Capturing texture inhomogeneity in medical volumetric images would also be beneficial in monitoring the development of diseases over longer periods of time to provide a better understanding of them.

8.2.2 Improving computation time

A further avenue of future work could be minimizing the computation time required for parameter estimation. While our proposed method achieves a significant reduction in the descriptor dimension by employing both binary encoding and spherical harmonics, it would be interesting to attempt to further increase the speed of the method. This could be done by improving the 3D GMRF parameter estimation or reducing the computation time of the interpolation technique used to interpolate neighbouring voxels.

8.2.3 Incorporating 3D GMRF models with deep learning-based methods

Various attempts have been made to integrate handcrafted feature-based methods with deep learning-based methods, such as in local binary convolutional neural networks (LBCNNs; Juefei-Xu et al. 2017), Local Binary Pattern Networks (LBPNets; Xi et al. 2016), and Gabor convolutional networks (GCNs; Luan et al. 2018), which incorporates Gabor filters into CNNs. The 3D GMRF model could be integrated with deep learning methods by fusing the texture features extracted by the 3D GMRF model with those learned by a deep learning method before classification or segmentation is performed (Xie, Jianpeng Zhang, Xia et al. 2018). Another approach would be to learn various variable settings of the 3D GMRF model through a learnable framework. This could lead to a significant reduction in the computational complexity of deep learning-based models (Juefei-Xu et al. 2017).

8.2.4 Comprehensive synthetic dataset

There are very few synthetic datasets available for volumetric texture analysis evaluation. Thus, generating more comprehensive synthetic datasets with clear ground truths is another promising avenue for future work. Creating such datasets for volumetric texture analysis involves generating a large number of labelled instances with complex 3D texture primitives and various levels of alterations (i.e., noise, rotation, translation, etc.). Further, it would be useful if the generated datasets included instances with inhomogeneous textures so that they are reliable for use in evaluating volumetric texture analysis methods.

References

A. P. Reeves A. M. Biancardi, s (2011). *The Lung Image Database Consortium (LIDC) Nodule Size Report.*

Abbasi, Solmaz and Farshad Tajeripour (2017). 'Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient'. In: *Neurocomputing* 219, pp. 526–535.

Akbari, Hamed and Baowei Fei (2012). 'Automatic 3D segmentation of the kidney in MR images using wavelet feature extraction and probability shape model'. In: *Medical Imaging 2012: Image Processing.* Vol. 8314. International Society for Optics and Photonics, p. 83143D.

Aldasoro, Constantino Carlos Reyes and Abhir Bhalerao (2007). 'Volumetric texture segmentation by discriminant feature selection and multiresolution classification'. In: *IEEE Transactions on Medical Imaging* 26.1, pp. 1–14.

Almaev, Timur R and Michel F Valstar (2013). 'Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition'. In: *2013 Humaine association conference on affective computing and intelligent interaction.* IEEE, pp. 356–361.

Andrearczyk, Vincent and Adrien Depeursinge (2018). 'Rotational 3D texture classification using group equivariant CNNs'. In: *arXiv preprint arXiv:1810.06889.*

Andrearczyk, Vincent, Julien Fageot, Valentin Oreiller, Xavier Montet and Adrien Depeursinge (2019). 'Exploring local rotation invariance in 3D CNNs with steerable filters'. In: *International Conference on Medical Imaging with Deep Learning*, pp. 15–26.

Andrearczyk, Vincent and Paul F Whelan (2018). 'Convolutional neural network on three orthogonal planes for dynamic texture classification'. In: *Pattern Recognition* 76, pp. 36–49.

Armato III, Samuel G, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I

Henschke, Eric A Hoffman et al. (2011). 'The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans'. In: *Medical physics* 38.2, pp. 915–931.

Banerjee, Jyotirmoy, Adriaan Moelker, Wiro J Niessen and Theo van Walsum (2012). '3D LBP-based rotationally invariant region description'. In: *Asian Conference on Computer Vision*. Springer, pp. 26–37.

Barbato, A1, T Frischer, CE Kuehni, D Snijders, I Azevedo, G Baktai, L Bartoloni, E Eber, A Escribano, E Haarman et al. (2009). 'Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children'. In: *European Respiratory Journal* 34.6, pp. 1264–1276.

Barra, Vincent and Jean-Yves Boire (2000). 'Tissue segmentation on MR images of the brain by possibilistic clustering on a 3D wavelet representation'. In: *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* 11.3, pp. 267–278.

Basu, Satrajit, Lawrence O Hall, Dmitry B Goldgof, Yuhua Gu, Virendra Kumar, Jung Choi, Robert J Gillies and Robert A Gatenby (2011). 'Developing a classifier model for lung tumors in CT-scan images'. In: *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, pp. 1306–1312.

El-Baz, Ayman, Manuel Casanova, G Gimel'farb, Meghan Mott, Andrew Switala, Eric Vanbogaert and Russ McCracken (2008). 'Dyslexia diagnostics by 3D texture analysis of cerebral white matter gyrifications'. In: *2008 19th International Conference on Pattern Recognition*. IEEE, pp. 1–4.

Belchi, Francisco, Mariam Pirashvili, Joy Conway, Michael Bennett, Ratko Djukanovic and Jacek Brodzki (2018). 'Lung topology characteristics in patients with chronic obstructive pulmonary disease'. In: *Scientific reports* 8.1, p. 5341.

Besag, Julian (1974). 'Spatial interaction and the statistical analysis of lattice systems'. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 36.2, pp. 192–225.

Bhatia, Parmeet S, Amit Kale and Zhigang Peng (2019). 'Volumetric texture modeling using dominant and discriminative binary patterns'. In: *Medical Imaging 2019: Image Processing*. Vol. 10949. International Society for Optics and Photonics, 109490H.

Björkström, A. (2001). In: *Stockholm University, Department of Mathematics*.

Blake, Andrew, Pushmeet Kohli and Carsten Rother (2011). *Markov random fields for vision and image processing*. The MIT Press.

Cao, Xiangyong, Feng Zhou, Lin Xu, Deyu Meng, Zongben Xu and John Paisley (2018). ‘Hyperspectral image classification with Markov random fields and a convolutional neural network’. In: *IEEE Transactions on Image Processing* 27.5, pp. 2354–2367.

Casanova, Dalcimar, João Batista Florindo, Mauricio Falvo and Odemir Martinez Bruno (2016). ‘Texture analysis using fractal descriptors estimated by the mutual interference of color channels’. In: *Information Sciences* 346, pp. 58–72.

Celli, Bartolome R, WATS MacNee, AATS Agusti, Antonio Anzueto, B Berg, A Sonia Buist, Peter MA Calverley, N Chavannes, T Dillard, B Fahy et al. (2004). ‘Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper’. In: *European Respiratory Journal* 23.6, pp. 932–946.

Chellappa, R. (1989). ‘Two-Dimensional Discrete Gaussian Markov Random Field Models for Image Processing and Analysis’. In: *Digital Image Processing Applications.* Ed. by Ying-Wei Lin and Ram Srinivasan. Vol. 1075. International Society for Optics and Photonics. SPIE, pp. 336–348. DOI: 10.1117/12.952662.

Chellappa, Rama and Shankar Chatterjee (1985). ‘Classification of textures using Gaussian Markov random fields’. In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33.4, pp. 959–963.

Chen, Hao, Qi Dou, Xi Wang, Jing Qin, Jack CY Cheng and Pheng-Ann Heng (2016). ‘3D fully convolutional networks for intervertebral disc localization and segmentation’. In: *International Conference on Medical Imaging and Augmented Reality.* Springer, pp. 375–382.

Chen, Hao, Qi Dou, Lequan Yu, Jing Qin and Pheng-Ann Heng (2018). ‘VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images’. In: *NeuroImage* 170, pp. 446–455.

Cheplygina, Veronika, Lauge Sørensen, David MJ Tax, Jesper Holst Pedersen, Marco Loog and Marleen de Bruijne (2014). ‘Classification of COPD with multiple instance learning’. In: *2014 22nd International Conference on pattern recognition.* IEEE, pp. 1508–1513.

Chopra, Satinder and Vladimir Alexeev (2006). ‘Applications of texture attribute analysis to 3D seismic data’. In: *The Leading Edge* 25.8, pp. 934–940.

Çiçek, Özgün, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox and Olaf Ronneberger (2016). ‘3D U-Net: learning dense volumetric segmentation from sparse annotation’. In: *International conference on medical image computing and computer-assisted intervention.* Springer, pp. 424–432.

Cid, Yashin Dicente, Henning Müller, Alexandra Platon, Pierre-Alexandre Poletti and Adrien Depeursinge (2017). '3D solid texture classification using locally-oriented wavelet transforms'. In: *IEEE Transactions on Image Processing* 26.4, pp. 1899–1910. ISSN: 1057-7149.

Ciompi, Francesco, Bartjan de Hoop, Sarah J van Riel, Kaman Chung, Ernst Th Scholten, Matthijs Oudkerk, Pim A de Jong, Mathias Prokop and Bram van Ginneken (2015). 'Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box'. In: *Medical image analysis* 26.1, pp. 195–202.

Cirujeda, Pol, Yashin Dicente Cid, Henning Müller, Daniel Rubin, Todd A Aguilera, Billy W Loo, Maximilian Diehn, Xavier Binefa and Adrien Depeursinge (2016). 'A 3-D Riesz-covariance texture model for prediction of nodule recurrence in lung CT'. In: *IEEE transactions on medical imaging* 35.12, pp. 2620–2630.

Cirujeda, Pol, Henning Müller, Daniel Rubin, Todd A Aguilera, Billy W Loo, Maximilian Diehn, Xavier Binefa and Adrien Depeursinge (2015). '3D Riesz-wavelet based Covariance descriptors for texture classification of lung nodule tissue in CT'. In: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, pp. 7909–7912.

Citraro, Leonardo, Sasan Mahmoodi, Angela Darekar and Brigitte Vollmer (2017). 'Extended three-dimensional rotation invariant local binary patterns'. In: *Image and Vision Computing* 62, pp. 8–18. ISSN: 0262-8856.

Cohen, Fernand S., Zhigang Fan and Maqbool A Patel (1991). 'Classification of rotated and scaled textured images using Gaussian Markov random field models'. In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 2, pp. 192–202.

Conci, Aura and Leonardo Hiss Monteiro (2000). 'Multifractal characterization of texture-based segmentation'. In: *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*. Vol. 1. IEEE, pp. 792–795.

Danciu, Marius, Mihaela Gordan, Camelia Florea and Aurel Vlaicu (2012). '3D DCT supervised segmentation applied on liver volumes'. In: *2012 35th International Conference on Telecommunications and Signal Processing (TSP)*. IEEE, pp. 779–783.

Danilla, Carolyne, Claudio Persello, Valentyn Tolpekin and John Ray Bergado (2017). 'Classification of multitemporal SAR images using convolutional neural networks and Markov random fields'. In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE, pp. 2231–2234.

De Nunzio, G, G Pastore, M Donativi, A Castellano and A Falini (2011). 'A CAD system for cerebral glioma based on texture features in DT-MR images'. In: *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 648, S100–S102.

Decramer, Marc, Wim Janssens and Marc Miravitlles (2012). 'Chronic obstructive pulmonary disease'. In: *The Lancet* 379.9823, pp. 1341–1351.

Deng, Huawu and David A Clausi (2004). 'Gaussian MRF rotation-invariant features for image classification'. In: *IEEE transactions on pattern analysis and machine intelligence* 26.7, pp. 951–955.

Depeursinge, Adrien, Julien Fageot, Vincent Andrearczyk, John Paul Ward and Michael Unser (2018). 'Rotation invariance and directional sensitivity: Spherical harmonics versus radiomics features'. In: *International Workshop on Machine Learning in Medical Imaging.* Springer, pp. 107–115.

Depeursinge, Adrien, Antonio Foncubierta-Rodriguez, Dimitri Van De Ville and Henning Müller (2014). 'Three-dimensional solid texture analysis in biomedical imaging: review and opportunities'. In: *Medical image analysis* 18.1, pp. 176–196.

Depeursinge, Adrien, Pedram Pad, Anne S Chin, Ann N Leung, Daniel L Rubin, Henning Müller and Michael Unser (2015). 'Optimized steerable wavelets for texture analysis of lung tissue in 3-D CT: Classification of usual interstitial pneumonia'. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI).* IEEE, pp. 403–406.

Depeursinge, Adrien, Daniel Sage, Asmâa Hidki, Alexandra Platon, Pierre-Alexandre Poletti, Michael Unser and Henning Muller (2007). 'Lung tissue classification using wavelet frames'. In: *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* IEEE, pp. 6259–6262.

Deserno, Markus (2004). 'How to generate equidistributed points on the surface of a sphere'. In: *If Polymerforshung (Ed.)* P. 99.

Dharmagunawardhana, Chathurika, Sasan Mahmoodi, Michael Bennett and Mahesan Niranjan (2014). 'Quantitative analysis of pulmonary emphysema using isotropic Gaussian Markov random fields'. In: *2014 International Conference on Computer Vision Theory and Applications (VISAPP).* Vol. 3. IEEE, pp. 44–53.

– (2016). 'Rotation invariant texture descriptors based on Gaussian Markov random fields for classification'. In: *Pattern Recognition Letters* 69, pp. 15–21.

El khadiri, I, A Chahi, Y Ruichek, R Touahni et al. (2018). 'Local directional ternary pattern: A new texture descriptor for texture classification'. In: *Computer vision and image understanding* 169, pp. 14–27.

Faust, Oliver, U Rajendra Acharya, Kristen M Meiburger, Filippo Molinari, Joel EW Koh, Chai Hong Yeong, Pailin Kongmebhol and Kwan Hoong Ng (2018). 'Comparative assessment of texture features for the identification of cancer in ultrasound images: a review'. In: *Biocybernetics and Biomedical Engineering* 38.2, pp. 275–296.

Fehr, Janis (2007). 'Rotational invariant uniform local binary patterns for full 3D volume texture analysis'. In: *Finnish signal processing symposium (FINSIG)*, pp. 1–7.

Fehr, Janis and Hans Burkhardt (2008). '3D rotation invariant local binary patterns'. In: *2008 19th International Conference on Pattern Recognition*. IEEE, pp. 1–4.

Fernández, Margarita, Adriana Mavilio and Manuel Tejera (2000). 'Texture segmentation of a 3D seismic section with wavelet transform and Gabor filters'. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. Vol. 3. IEEE, pp. 354–357.

Florindo, Joao Batista, André Ricardo Backes, Mário de Castro and Odemir Martinez Bruno (2012). 'A comparative study on multiscale fractal dimension descriptors'. In: *Pattern Recognition Letters* 33.6, pp. 798–806.

Friedman, J., T. Hastie and R. Tibshirani (2001). In: *The Elements of Statistical Learning*. Vol. vol. 1. Springer series in statistics New York.

Funke, Christina M, Leon A Gatys, Alexander S Ecker and Matthias Bethge (2017). 'Synthesising dynamic textures using convolutional neural networks'. In: *arXiv preprint arXiv:1702.07006*.

Galloway, Mary M (1974). 'Texture analysis using grey level run lengths'. In: *STIN* 75, p. 18555.

Gangeh, Mehrdad J, Lauge Sørensen, Saher B Shaker, Mohamed S Kamel, Marleen De Bruijne and Marco Loog (2010). 'A texton-based approach for the classification of lung parenchyma in CT images'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 595–602.

Gao, Dengliang (2003). 'Volume texture extraction for 3D seismic visualization and interpretation'. In: *Geophysics* 68.4, pp. 1294–1302.

Gao, Xiaohong, Yu Qian, Rui Hui, Martin Loomes, Richard Comley, Balbir Barn, A Chapman and J Rix (2010). 'Texture-based 3D image retrieval for medical applications'. In: *IADIS International Conference e-Health*, pp. 101–108.

Genschel, Ulrike and William Q Meeker (2010). 'A comparison of maximum likelihood and median-rank regression for Weibull estimation'. In: *Quality Engineering* 22.4, pp. 236–255.

Gonçalves, Wesley Nunes, Bruno Brandoli Machado and Odemir Martinez Bruno (2012). 'Spatiotemporal gabor filters: a new method for dynamic texture recognition'. In: *arXiv preprint arXiv:1201.3612*.

Green, Robin (2003). 'Spherical harmonic lighting: The gritty details'. In: *Archives of the Game Developers Conference*. Vol. 56, p. 4.

Griffiths, David and Jan Boehm (2019). 'A Review on deep learning techniques for 3D sensed data classification'. In: *Remote Sensing* 11.12, p. 1499.

Hackx, Maxime, Alexander A Bankier and Pierre Alain Gevenois (2012). 'Chronic obstructive pulmonary disease: CT quantification of airways disease'. In: *Radiology* 265.1, pp. 34–48.

Hafner, M, A Gangl, M Liedlgruber, A Uhl, A Vécsei and F Wrba (2009). 'Combining gaussian Markov random fields with the discrete-wavelet transform for endoscopic image classification'. In: *2009 16th international conference on digital signal processing*. IEEE, pp. 1–6.

Han, Fangfang, Huafeng Wang, Guopeng Zhang, Hao Han, Bowen Song, Lihong Li, William Moore, Hongbing Lu, Hong Zhao and Zhengrong Liang (2015). 'Texture feature analysis for computer-aided diagnosis on pulmonary nodules'. In: *Journal of digital imaging* 28.1, pp. 99–115.

Haralick, Robert M (1979). 'Statistical and structural approaches to texture'. In: *Proceedings of the IEEE* 67.5, pp. 786–804.

Haralick, Robert M, Karthikeyan Shanmugam and Its' Hak Dinstein (1973). 'Textural features for image classification'. In: *IEEE Transactions on systems, man, and cybernetics* 6, pp. 610–621.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren and Jian Sun (2016). 'Deep residual learning for image recognition'. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

Hett, Kilian, Vinh-Thong Ta, José V Manjón, Pierrick Coupé, Alzheimer's Disease Neuroimaging Initiative et al. (2018). 'Adaptive fusion of texture-based grading for Alzheimer's disease classification'. In: *Computerized Medical Imaging and Graphics* 70, pp. 8–16.

Horn, Berthold KP and Brian G Schunck (1981). 'Determining optical flow'. In: *Artificial intelligence* 17.1-3, pp. 185–203.

Hussein, Sarfaraz, Kunlin Cao, Qi Song and Ulas Bagci (2017). 'Risk stratification of lung nodules using 3d cnn-based multi-task learning'. In: *International conference on information processing in medical imaging*. Springer, pp. 249–260.

Ikegami, Koji, Showbu Sato, Kenji Nakamura, Lawrence E Ostrowski and Mitsutoshi Setou (2010). ‘Tubulin polyglutamylation is essential for airway ciliary function through the regulation of beating asymmetry’. In: *Proceedings of the National Academy of Sciences* 107.23, pp. 10490–10495.

Jafari-Khouzani, Kourosh, Hamid Soltanian-Zadeh, Kost Elisevich and Suresh Patel (2004). ‘Comparison of 2D and 3D wavelet features for TLE lateralization’. In: *Medical Imaging 2004: Physiology, Function, and Structure from Medical Images.* Vol. 5369. International Society for Optics and Photonics, pp. 593–602.

Jain, Anil K and Farshid Farrokhnia (1990). ‘Unsupervised texture segmentation using Gabor filters’. In: *1990 IEEE international conference on systems, man, and cybernetics conference proceedings.* IEEE, pp. 14–19.

Jalalian, Afsaneh, Syamsiah Mashohor, Rozi Mahmud, Babak Karasfi, M Iqbal Saripan and Abdul Rahman Ramli (2017). ‘Computer-assisted diagnosis system for breast cancer in computed tomography laser mammography (CTLM)’. In: *Journal of digital imaging* 30.6, pp. 796–811.

Juefei-Xu, Felix, Vishnu Naresh Boddeti and Marios Savvides (2017). ‘Local binary convolutional neural networks’. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 19–28.

Julesz, Bela (1981). ‘Textons, the elements of texture perception, and their interactions’. In: *Nature* 290.5802, pp. 91–97.

Al-Kadi, Omar S (2017). ‘Fractals for biomedical texture analysis’. In: *Biomedical Texture Analysis.* Elsevier, pp. 131–161.

Kashyap, Rangasami L, Rama Chellappa and Alireza Khotanzad (1982). ‘Texture classification using features derived from random field models’. In: *Pattern Recognition Letters* 1.1, pp. 43–50.

Kazhdan, Michael, Thomas Funkhouser and Szymon Rusinkiewicz (2003). ‘Rotation invariant spherical harmonic representation of 3D shape descriptors’. In: *Symposium on geometry processing.* Vol. 6, pp. 156–164.

Kim, Woojae, Tae Hwa Han, Hyun Jun Kim, Man Young Park, Ku Sang Kim and Rae Woong Park (2011). ‘An automated measurement of ciliary beating frequency using a combined optical flow and peak detection’. In: *Healthcare informatics research* 17.2, pp. 111–119.

Korfiatis, Panayiotis D, Anna N Karahaliou, Alexandra D Kazantzi, Cristina Kalogeropoulou and Lena I Costaridou (2009). 'Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT'. In: *IEEE transactions on information technology in biomedicine* 14.3, pp. 675–680.

Kovalev, Vassili A, Frithjof Kruggel, H-J Gertz and D Yves von Cramon (2001). 'Three-dimensional texture analysis of MRI brain datasets'. In: *IEEE transactions on medical imaging* 20.5, pp. 424–433.

Kovalev, Vassili A, Maria Petrou and Yaroslav S Bondar (1999). 'Texture anisotropy in 3-D images'. In: *IEEE Transactions on Image Processing* 8.3, pp. 346–360.

Krizhevsky, Alex, Ilya Sutskever and Geoffrey E Hinton (2012). 'Imagenet classification with deep convolutional neural networks'. In: *Advances in neural information processing systems*, pp. 1097–1105.

Kumar, Ashnil, Jinman Kim, David Lyndon, Michael Fulham and Dagan Feng (2016). 'An ensemble of fine-tuned convolutional neural networks for medical image classification'. In: *IEEE journal of biomedical and health informatics* 21.1, pp. 31–40.

Kurani, Arati S, Dong-Hui Xu, Jacob Furst and Daniela Stan Raicu (2004). 'Co-occurrence matrices for volumetric data'. In: *Heart* 27, p. 25.

Leung, Thomas and Jitendra Malik (2001). 'Representing and recognizing the visual appearance of materials using three-dimensional textons'. In: *International journal of computer vision* 43.1, pp. 29–44.

Li, Stan Z (2009). *Markov random field modeling in image analysis*. Springer Science & Business Media.

Li, Xin, LZ Tong, XX Zhou and Xu Wang (2011). 'Classification of 3D texture features based on MR image in discrimination of Alzheimer disease and mild cognitive impairment from normal controls'. In: *Chinese Journal of Medical Imaging Technology* 27.5, pp. 1047–1051.

Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken and Clara I Sánchez (2017). 'A survey on deep learning in medical image analysis'. In: *Medical image analysis* 42, pp. 60–88.

Liu, Sidong, Weidong Cai, Lingfeng Wen and D Feng (2011). 'Volumetric congruent local binary patterns for 3D neurological image retrieval'. In: *International conference on image and vision computing New Zealand (ICVNZ)*, pp. 272–6.

Liu, Xiuwen and DeLiang Wang (2003). 'Texture classification using spectral histograms'. In: *IEEE transactions on image processing* 12.6, pp. 661–670.

Lopes, Renaud, P Dubois, Imen Bhouri, Mohamed Hedi Bedoui, Salah Maouche and Nacim Betrouni (2011). ‘Local fractal and multifractal features for volumic texture characterization’. In: *Pattern Recognition* 44.8, pp. 1690–1697.

Lowe, David G (2004). ‘Distinctive image features from scale-invariant keypoints’. In: *International journal of computer vision* 60.2, pp. 91–110.

Luan, Shangzhen, Chen Chen, Baochang Zhang, Jungong Han and Jianzhuang Liu (2018). ‘Gabor convolutional networks’. In: *IEEE Transactions on Image Processing* 27.9, pp. 4357–4366.

Lucas, Jane S, Andrea Burgess, Hannah M Mitchison, Eduardo Moya, Michael Williamson and Claire Hogg (2014). ‘Diagnosis and management of primary ciliary dyskinesia’. In: *Archives of disease in childhood* 99.9, pp. 850–856.

Luengo, Imanol, Mark Basham and Andrew P French (2016). ‘Fast global interactive volume segmentation with regional supervoxel descriptors’. In: *Medical Imaging 2016: Image Processing*. Vol. 9784. International Society for Optics and Photonics, p. 97842D.

Luengo, Imanol, Michele C Darrow, Matthew C Spink, Ying Sun, Wei Dai, Cynthia Y He, Wah Chiu, Tony Pridmore, Alun W Ashton, Elizabeth MH Duke et al. (2017). ‘SuRVoS: Super-region volume segmentation workbench’. In: *Journal of Structural Biology* 198.1, pp. 43–53.

Lynch, David A (2014). ‘Progress in imaging COPD, 2004-2014’. In: *Chronic Obstructive Pulmonary Diseases: Journal of the COPD Foundation* 1.1, p. 73.

Maani, Rouzbeh, Sanjay Kalra and Yee-Hong Yang (2014). ‘Robust volumetric texture classification of magnetic resonance images of the brain using local frequency descriptor’. In: *IEEE Transactions on Image Processing* 23.10, pp. 4625–4636.

Madabhushi, Anant, Michael Feldman, Dimitris Metaxas, Deborah Chute and John Tomaszewski (2003). ‘A novel stochastic combination of 3D texture features for automated segmentation of prostatic adenocarcinoma from high resolution MRI’. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 581–591.

Mahmoud-Ghoneim, Doaa, Grégoire Toussaint, Jean-Marc Constans and D Jacques (2003). ‘Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas’. In: *Magnetic resonance imaging* 21.9, pp. 983–987.

Manjunath, Bangalore S and Wei-Ying Ma (1996). ‘Texture features for browsing and retrieval of image data’. In: *IEEE Transactions on pattern analysis and machine intelligence* 18.8, pp. 837–842.

Mariolis, I, Panagiotis Korfiatis, C Kalogeropoulou, D Daoussis, T Petsas and L Costaridou (2010). 'Computer aided diagnosis of diffuse lung disease in multi-detector CT–selecting 3D texture features'. In: *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010*. Springer, pp. 208–211.

Mascalchi, Mario, Gianna Camiciottoli and Stefano Diciotti (2017). 'Lung densitometry: why, how and when'. In: *Journal of thoracic disease* 9.9, p. 3319.

Mishra, Ashwini Kumar, Desok Kim and Ilma Andayana (2011). 'Development of three dimensional binary patterns for local bone structure analysis'. In: *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. IEEE, pp. 1006–1008.

Mondol, Nazmul Haque (2010). 'Seismic exploration'. In: *Petroleum Geoscience*. Springer, pp. 375–402.

Morgado, Pedro, Margarida Silveira and Jorge S Marques (2013). 'Diagnosis of Alzheimer's disease using 3D local binary patterns'. In: *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 1.1, pp. 2–12.

Morillas, Hilda N, Maimoona Zariwala and Michael R Knowles (2007). 'Genetic causes of bronchiectasis: primary ciliary dyskinesia'. In: *Respiration* 74.3, pp. 252–263.

Muer, Nestor L, Catherine A Staples, Roberta R Miller and Raja T Abboud (1988). '"Density mask": an objective method to quantitate emphysema using computed tomography'. In: *Chest* 94.4, pp. 782–787. ISSN: 0012-3692.

Müller, Nestor L, Catherine A Staples, Roberta R Miller and Raja T Abboud (1988). '"Density mask": an objective method to quantitate emphysema using computed tomography'. In: *Chest* 94.4, pp. 782–787.

Needham, Chris J and Roger D Boyle (2003). 'Performance evaluation metrics and statistics for positional tracker evaluation'. In: *International Conference on Computer Vision Systems*. Springer, pp. 278–289.

Nixon, Mark and Alberto Aguado (2008). *Feature extraction and image processing for computer vision*. Academic Press.

Noone, Peadar G, William D Bennett, Jeff A Regnis, Kirby L Zeman, Johnny L Carson, Malcolm King, Richard C Boucher and Michael R Knowles (1999). 'Effect of aerosolized uridine-5 -triphosphate on airway clearance with cough in patients with primary ciliary dyskinesia'. In: *American journal of respiratory and critical care medicine* 160.1, pp. 144–149.

Ojala, Timo, Matti Pietikäinen and David Harwood (1996). ‘A comparative study of texture measures with classification based on featured distributions’. In: *Pattern recognition* 29.1, pp. 51–59.

Ojala, Timo, Matti Pietikäinen and Topi Maenpaa (2002). ‘Multiresolution gray-scale and rotation invariant texture classification with local binary patterns’. In: *IEEE Transactions on pattern analysis and machine intelligence* 24.7, pp. 971–987.

Omran, Heymut and Niki T Loges (2009). ‘Immunofluorescence staining of ciliated respiratory epithelial cells’. In: *Methods in cell biology.* Vol. 91. Elsevier, pp. 123–133.

Othmen, Elmoez Ben, Mounir Sayadi and Farhat Fniaech (2013). ‘3D gray level co-occurrence matrices for volumetric texture classification’. In: *3rd International Conference on Systems and Control.* IEEE, pp. 833–837.

Päivärinta, Juhani, Esa Rahtu and Janne Heikkilä (2011). ‘Volume local phase quantization for blur-insensitive dynamic texture classification’. In: *Scandinavian Conference on Image Analysis.* Springer, pp. 360–369.

Papon, JF, A Coste, F Roudot-Thoraval, M Boucherat, G Roger, A Tamalet, AM Vojtek, S Amselem and E Escudier (2010). ‘A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia’. In: *European Respiratory Journal* 35.5, pp. 1057–1063.

Park, Yang Shin, Joon Beom Seo, Namkug Kim, Eun Jin Chae, Yeon Mok Oh, Sang Do Lee, Youngjoo Lee and Suk-Ho Kang (2008). ‘Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: comparison with density-based quantification and correlation with pulmonary function test’. In: *Investigative radiology* 43.6, pp. 395–402.

Patel, Daniel, Stefan Bruckner, Ivan Viola and Eduard M Gröller (2010). ‘Seismic volume visualization for horizon extraction’. In: *2010 IEEE Pacific Visualization Symposium (PacificVis).* IEEE, pp. 73–80.

Paulhac, Ludovic, Pascal Makris and Jean-Yves Ramel (2008). ‘Comparison between 2D and 3D local binary pattern methods for characterisation of three-dimensional textures’. In: *International Conference Image Analysis and Recognition.* Springer, pp. 670–679.

Paulhac, Ludovic, Pascal Makris, Jean-Yves Ramel et al. (2009). ‘A solid texture database for segmentation and classification experiments.’ In: *VISAPP (2)*, pp. 135–141.

Paulhac, Ludovic, Pascal Makris, Jean-Yves Ramel and Jean-Marc Gregoire (2015). ‘A framework of perceptual features for the characterisation of 3D textured images’. In: *Signal, Image and Video Processing* 9.2, pp. 305–329. ISSN: 1863-1703.

Pena, Isabel Pino, Veronika Cheplygina, Sofia Paschaloudi, Morten Vuust, Jesper Carl, Ulla Møller Weinreich, Lasse Riis Østergaard and Marleen de Bruijne (2018). 'Automatic emphysema detection using weakly labeled HRCT lung images'. In: *PloS one* 13.10.

Petrou, Maria and Pedro Garcia Sevilla (2006). *Image processing: dealing with texture.* Wiley.

Puybareau, E, Hugues Talbot, Gabriel Pelle, Bruno Louis, J-F Papon, André Coste and Laurent Najman (2015). 'A regionalized automated measurement of ciliary beating frequency'. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI).* IEEE, pp. 528–531.

Qaseem, Amir, Timothy J Wilt, Steven E Weinberger, Nicola A Hanania, Gerard Criner, Thys van der Molen, Darcy D Marciniuk, Tom Denberg, Holger Schunemann, Wisia Wedzicha et al. (2011). 'Diagnosis and management of stable chronic obstructive pulmonary disease: a clinical practice guideline update from the American College of Physicians, American College of Chest Physicians, American Thoracic Society, and European Respiratory Society'. In: *Annals of internal medicine* 155.3, pp. 179–191.

Quellec, Gwénolé, Kyungmoo Lee, Martin Dolejsi, Mona K Garvin, Michael D Abramoff and Milan Sonka (2010). 'Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in SD-OCT of the macula'. In: *IEEE transactions on medical imaging* 29.6, pp. 1321–1330.

Quinn, Shannon P, Maliha J Zahid, John R Durkin, Richard J Francis, Cecilia W Lo and S Chakra Chennubhotla (2015). 'Automated identification of abnormal respiratory ciliary motion in nasal biopsies'. In: *Science translational medicine* 7.299, 299ra124–299ra124.

Ranguelova, Elena and Anthony Quinn (1999). 'Analysis and synthesis of three-dimensional Gaussian Markov random fields'. In: *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348).* Vol. 3. IEEE, pp. 430–434.

Rellier, Guillaume, Xavier Descombes, Frederic Falzon and Josiane Zerubia (2004). 'Texture feature analysis using a Gauss-Markov model in hyperspectral image classification'. In: *IEEE Transactions on Geoscience and Remote Sensing* 42.7, pp. 1543–1551.

Rue, Havard and Leonhard Held (2005). *Gaussian Markov random fields: theory and applications.* CRC press.

Rumman, Nisreen, Claire Jackson, Samuel Collins, Patricia Goggin, Janice Coles and Jane S Lucas (2017). 'Diagnosis of primary ciliary dyskinesia: potential options for resource-limited countries'. In: *European Respiratory Review* 26.143.

Sarwar, Syed Shakib, Priyadarshini Panda and Kaushik Roy (2017). ‘Gabor filter assisted energy efficient fast learning convolutional neural networks’. In: *2017 IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED)*. IEEE, pp. 1–6.

Shafer, Christina M, Victoria L Seewaldt and Joseph Y Lo (2011). ‘Validation of a 3D hidden-Markov model for breast tissue segmentation and density estimation from MR and tomosynthesis images’. In: *Proceedings of the 2011 Biomedical Sciences and Engineering Conference: Image Informatics and Analytics in Biomedicine*. IEEE, pp. 1–4.

Shamsheyeva, Alena and Arcot Sowmya (2004). ‘The anisotropic Gaussian kernel for SVM classification of HRCT images of the lung’. In: *Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004*. IEEE, pp. 439–444.

Shao, Rui, Xiangyuan Lan and Pong C Yuen (2017). ‘Deep convolutional dynamic texture learning with adaptive channel-discriminability for 3D mask face anti-spoofing’. In: *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, pp. 748–755.

Shapiro, Adam J, Maimoona A Zariwala, Thomas Ferkol, Stephanie D Davis, Scott D Sagel, Sharon D Dell, Margaret Rosenfeld, Kenneth N Olivier, Carlos Milla, Sam J Daniel et al. (2016). ‘Diagnosis, monitoring, and treatment of primary ciliary dyskinesia: PCD foundation consensus recommendations based on state of the art review’. In: *Pediatric pulmonology* 51.2, pp. 115–132.

Shen, Wei, Mu Zhou, Feng Yang, Dongdong Yu, Di Dong, Caiyun Yang, Yali Zang and Jie Tian (2017). ‘Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification’. In: *Pattern Recognition* 61, pp. 663–673.

Shoemark, A, M Dixon, B Corrin and A Dewar (2012). ‘Twenty-year review of quantitative transmission electron microscopy for the diagnosis of primary ciliary dyskinesia’. In: *Journal of clinical pathology* 65.3, pp. 267–271.

Sluimer, Ingrid C, Paul F van Waes, Max A Viergever and Bram van Ginneken (2003). ‘Computer-aided diagnosis in high resolution CT of the lungs’. In: *Medical physics* 30.12, pp. 3081–3090.

Sorensen, Lauge, Mads Nielsen, Pechin Lo, Haseem Ashraf, Jesper H Pedersen and Marleen De Bruijne (2012). ‘Texture-based analysis of COPD: a data-driven approach’. In: *IEEE transactions on medical imaging* 31.1, pp. 70–78.

Sorensen, Lauge, Saher B Shaker and Marleen De Bruijne (2010). ‘Quantitative analysis of pulmonary emphysema using local binary patterns’. In: *IEEE transactions on medical imaging* 29.2, pp. 559–569.

Stannard, Wendy A, Mark A Chilvers, Andrew R Rutman, Chris D Williams and Chris O’Callaghan (2010). ‘Diagnostic testing of patients suspected of primary ciliary dyskinesia’. In: *American journal of respiratory and critical care medicine* 181.4, pp. 307–314.

Stoel, Berend C and Jan Stolk (2004). ‘Optimization and standardization of lung densitometry in the assessment of pulmonary emphysema’. In: *Investigative radiology* 39.11, pp. 681–688. ISSN: 0020-9996.

Tajbakhsh, Nima, Jae Y Shin, Suryakanth R Gurudu, R Todd Hurst, Christopher B Kendall, Michael B Gotway and Jianming Liang (2016). ‘Convolutional neural networks for medical image analysis: Full training or fine tuning?’ In: *IEEE transactions on medical imaging* 35.5, pp. 1299–1312.

Thibault, Guillaume, Alina Tudorica, Aneela Afzal, Stephen YC Chui, Arpana Naik, Megan L Troxell, Kathleen A Kemmer, Karen Y Oh, Nicole Roy, Neda Jafarian et al. (2017). ‘DCE-MRI texture features for early prediction of breast cancer therapy response’. In: *Tomography* 3.1, p. 23.

Toro, Oscar Alfonso Jiménez del, Antonio Foncubierta–Rodriguez, María Isabel Vargas Gómez, Henning Müller and Adrien Depeursinge (2013). ‘Epileptogenic lesion quantification in MRI using contralateral 3D texture comparisons’. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, pp. 353–360.

Tuceryan, Mihran and Anil K Jain (1993). ‘Texture analysis’. In: *Handbook of pattern recognition and computer vision.* World Scientific, pp. 235–276.

Upadhyay, Sanat, Manos Papadakis, Saurabh Jain, Gregory Gladish, Ioannis A Kakadiaris and Robert Azencott (2012). ‘Semi-automatic discrimination of normal tissue and liver cancer lesions in contrast enhanced X-ray CT-scans’. In: *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging.* Springer, pp. 158–167.

Uppaluri, Renuka, Eric A Hoffman, Milan Sonka, Patrick G Hartley, Gary W Hunninghake and Geoffrey McLennan (1999). ‘Computer recognition of regional lung disease patterns’. In: *American journal of respiratory and critical care medicine* 160.2, pp. 648–654.

Uppaluri, Renuka, Eric A Hoffman, Milan Sonka, Gary W Hunninghake and Geoffrey McLennan (1999). ‘Interstitial lung disease: a quantitative study using the adaptive multiple feature method'. In: *American journal of respiratory and critical care medicine* 159.2, pp. 519–525. ISSN: 1535-4970.

Uppaluri, Renuka, Theophano Mitsa, Milan Sonka, Eric A Hoffman and Geoffrey McLennan (1997). 'Quantification of pulmonary emphysema from lung computed tomography images'. In: *American journal of respiratory and critical care medicine* 156.1, pp. 248–254.

Vasconcelos, Verónica, José Silvestre Silva, Luís Marques and João Barroso (2010). 'Statistical textural features for classification of lung emphysema in CT images: A comparative study'. In: *5th Iberian Conference on Information Systems and Technologies*. IEEE, pp. 1–5.

Vogelmeier, Claus F, Gerard J Criner, Fernando J Martinez, Antonio Anzueto, Peter J Barnes, Jean Bourbeau, Bartolome R Celli, Rongchang Chen, Marc Decramer, Leonardo M Fabbri et al. (2017). 'Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 report. GOLD executive summary'. In: *American journal of respiratory and critical care medicine* 195.5, pp. 557–582.

Wang, Chaohui, Nikos Komodakis and Nikos Paragios (2013). 'Markov random field modeling, inference & learning in computer vision & image understanding: A survey'. In: *Computer Vision and Image Understanding* 117.11, pp. 1610–1627.

Wang, Liping and Reyer Zwiggelaar (2017). '3d texton based prostate cancer detection using multiparametric magnetic resonance imaging'. In: *Annual conference on medical image understanding and analysis*. Springer, pp. 309–319.

Wang, Zhimin, Suicheng Gu, Joseph K Leader, Shinjini Kundu, John R Tedrow, Frank C Sciurba, David Gur, Jill M Siegfried and Jiantao Pu (2013). 'Optimal threshold in CT quantification of emphysema'. In: *European radiology* 23.4, pp. 975–984.

Waugh, SA, CA Purdie, LB Jordan, S Vinnicombe, RA Lerski, P Martin and AM Thompson (2016). 'Magnetic resonance imaging texture analysis classification of primary breast cancer'. In: *European radiology* 26.2, pp. 322–330.

Winkels, Marysia and Taco S Cohen (2018). '3d g-cnns for pulmonary nodule detection'. In: *arXiv preprint arXiv:1804.04656*.

– (2019). 'Pulmonary nodule detection in CT scans with equivariant CNNs'. In: *Medical image analysis* 55, pp. 15–26.

'Data Mining: Practical Machine Learning Tools and Techniques' (2017). In: *Data Mining (Fourth Edition)*. Ed. by Ian H. Witten, Eibe Frank, Mark A. Hall and Christopher J. Pal. Fourth Edition. Morgan Kaufmann, pp. 335–416. ISBN: 978-0-12-804291-5.

Worrall, Daniel E, Stephan J Garbin, Daniyar Turmukhambetov and Gabriel J Brostow (2017). 'Harmonic networks: Deep translation and rotation equivariance'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5028–5037.

Xi, Meng, Liang Chen, Desanka Polajnar and Weiyang Tong (2016). 'Local binary pattern network: a deep learning approach for face recognition'. In: *2016 IEEE international conference on Image processing (ICIP)*. IEEE, pp. 3224–3228.

Xie, Yutong, Yong Xia, Jianpeng Zhang, Yang Song, Dagan Feng, Michael Fulham and Weidong Cai (2018). 'Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT'. In: *IEEE transactions on medical imaging* 38.4, pp. 991–1004.

Xie, Yutong, Jianpeng Zhang and Yong Xia (2019). 'Semi-supervised adversarial model for benign–malignant lung nodule classification on chest CT'. In: *Medical image analysis* 57, pp. 237–248.

Xie, Yutong, Jianpeng Zhang, Yong Xia, Michael Fulham and Yanning Zhang (2018). 'Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT'. In: *Information Fusion* 42, pp. 102–110.

Xu, Dong-Hui, Arati S Kurani, Jacob D Furst and Daniela S Raicu (2004). 'Run-length encoding for volumetric texture'. In: *Heart* 27.25, pp. 452–458.

Xu, Ye, Milan Sonka, Geoffrey McLennan, Junfeng Guo and Eric Hoffman (2005). 'Sensitivity and specificity of 3-D texture analysis of lung parenchyma is better than 2-D for discrimination of lung pathology in stage 0 COPD'. In: *Medical Imaging 2005: Physiology, Function, and Structure from Medical Images*. Vol. 5746. International Society for Optics and Photonics, pp. 474–485.

Xu, Yong, Sibin Huang, Hui Ji and Cornelia FermüLler (2012). 'Scale-space texture description on SIFT-like textons'. In: *Computer Vision and Image Understanding* 116.9, pp. 999–1013.

Yan, Xingjian, Jianing Pang, Hang Qi, Yixin Zhu, Chunxue Bai, Xin Geng, Mina Liu, Demetri Terzopoulos and Xiaowei Ding (2016). 'Classification of lung nodule malignancy risk on computed tomography images using convolutional neural network: A comparison between 2d and 3d strategies'. In: *Asian Conference on Computer Vision*. Springer, pp. 91–101.

Yang, Jie, Elsa D Angelini, Pallavi P Balte, Eric A Hoffman, John HM Austin, Benjamin M Smith, Jingkuan Song, R Graham Barr and Andrew F Laine (2017). 'Unsupervised discovery of spatially-informed lung texture patterns for pulmonary emphysema: The

MESA COPD study'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 116–124.

Yang, Yi and Shawn Newsam (2008). 'Comparing SIFT descriptors and Gabor texture features for classification of remote sensed imagery'. In: *2008 15th IEEE international conference on image processing*. IEEE, pp. 1852–1855.

Yao, Hu, Li Chuyi, Hu Dan and Yu Weiyu (2016). 'Gabor feature based convolutional neural network for object recognition in natural scene'. In: *2016 3rd International Conference on Information Science and Control Engineering (ICISCE)*. IEEE, pp. 386–390.

Yenugu, Malleswar, Kurt J Marfurt and Shane Matson (2010). 'Seismic texture analysis for reservoir prediction and characterization'. In: *The Leading Edge* 29.9, pp. 1116–1121.

Zhan, Yiqiang and Dinggang Shen (2003). 'Automated segmentation of 3D US prostate images using statistical texture-based matching method'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, pp. 688–696.

Zhang, Jianguo, Marcin Marszałek, Svetlana Lazebnik and Cordelia Schmid (2007). 'Local features and kernels for classification of texture and object categories: A comprehensive study'. In: *International journal of computer vision* 73.2, pp. 213–238.

Zhang, Jing, Chunshui Yu, Guilian Jiang, Weifang Liu and Longzheng Tong (2012). '3D texture analysis on MRI images of Alzheimer's disease'. In: *Brain imaging and behavior* 6.1, pp. 61–69.

Zhao, Guoying and Matti Pietikainen (2007). 'Dynamic texture recognition using local binary patterns with an application to facial expressions'. In: *IEEE transactions on pattern analysis and machine intelligence* 29.6, pp. 915–928.

Zhao, Xiaochao, Yaping Lin, Li Liu, Janne Heikkilä and Wenming Zheng (2019). 'Dynamic Texture Classification Using Unsupervised 3D Filter Learning and Local Binary Encoding'. In: *IEEE Transactions on Multimedia* 21.7, pp. 1694–1708.

Zhao, Yindi, Liangpei Zhang, Pingxiang Li and Bo Huang (2007). 'Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features'. In: *IEEE Transactions on Geoscience and Remote Sensing* 45.5, pp. 1458–1468.