UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Image Texture Analysis based on Gaussian Markov Random Fields

by

Chathurika Dharmagunawardhana

Thesis for the degree of Doctor of Philosophy

November 2014

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
Electronics and Computer Science

Doctor of Philosophy

IMAGE TEXTURE ANALYSIS BASED ON GAUSSIAN MARKOV RANDOM FIELDS

by Chathurika Dharmagunawardhana

Texture analysis is one of the key techniques of image understanding and processing with widespread applications from low level image segmentation to high level object recognition. Gaussian Markov random field (GMRF) is a particular model based texture feature extraction scheme which uses model parameters as texture features. In this thesis a novel robust texture descriptor based on GMRF is proposed specially for texture segmentation and classification. For these tasks, descriptive features are more favourable relative to the generative features. Therefore, in order to achieve more descriptive features, with the GMRFs, a localized parameter estimation technique is introduced here. The issues arising in the localized parameter estimation process, due to the associated small sample size, are addressed by applying Tikhonov regularization and an estimation window size selection criterion. The localized parameter estimation process proposed here can overcome the problem of parameter smoothing that occurs in traditional GMRF parameter estimation. Such a parameter smoothing disregards some important structural and statistical information for texture discrimination. The normalized distributions of local parameter estimates are proposed as the new texture features and are named as Local Parameter Histogram (LPH) descriptors. Two new rotation invariant texture descriptors based on LPH features are also introduced, namely Rotation Invariant LPH (RI-LPH) and Isotropic LPH (I-LPH) descriptors. The segmentation and classification results on large texture datasets demonstrate that these descriptors significantly improve the performance of traditional GMRF features and also demonstrate better performance in comparison with the state-of-the-art texture descriptors. Satisfactory natural image segmentation is also carried out based on the novel features. Furthermore, proposed features are employed in a real world medical application involving tissue recognition for emphysema, a critical lung disease causing lung tissue destruction. Features extracted from High Resolution Computed Tomography (HRCT) data are used in effective tissue recognition and pathology distribution diagnosis. Moreover, preliminary work on a Bayesian framework for integrating prior knowledge into the parameter estimation process is undertaken to introduce further improved texture features.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Chathurika Dharmagunawardhana , declare that the thesis entitled *Image Texture Analysis based on 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;
- none of this work has been published before submission

Signed:.....

Date:10/11/2014.....

Acknowledgements

First and foremost, I would like to express my special thanks to my main supervisor, Dr Sasan Mahmoodi, for the continuous support and guidance he has given me throughout this research. I owe many thanks to Dr Michael Bennett for sharing his kind views and opinions with us, always keeping a very positive attitude. I am also thankful to Professor Mahesan Niranjan for the support he has given me despite of his busy schedule. I would like to thank my internal examiner, Professor Mark Nixon, for his encouraging and constructive feedbacks and my external examiner, Professor Josef Kittler from University of Surrey for his valuable comments and feedbacks.

I thank School of Electronics and Computer Science of University of Southampton for funding this research. I like to acknowledge the help from my friends and colleagues for all their support and for making my PhD life interesting as well as challenging.

Finally and specially, I would like to thank my loving husband and our families for supporting and comforting me during the entire period of the doctoral study. Without their trust and continuous motivation this thesis would not have been possible.

To my family...

Chapter 1

Context and Contributions

1.1 Context

Texture is a surface characteristic of an object or a region. It exhibits important information about the structural arrangements and distributions of primitive elements or patterns which relate to the formation of the surface appearance. Texture plays a vital role in human perception of visual objects and scenic regions, together with other visual cues such as colour, brightness and form (Tuceryan and Jain, 1998).

Texture analysis has its own significance in computer vision with widespread applications in many fields including medical image processing, remote sensing, document processing, defect detection, image retrieval, object recognition and computer graphic generation. In image texture analysis a particular texture is characterized in terms of texture features. Texture feature extraction mainly aims at formulating effective discriminative texture descriptors. The texture features extracted from the image textures are subsequently used in texture segmentation, classification or synthesis (Petrou and Sevilla, 2006; Tuceryan and Jain, 1998).

Texture feature extraction has been extensively studied in recent years and a large number of texture feature extraction techniques have been developed (Nixon and Aguado, 2008; Varma and Zisserman, 2009). These methods can be roughly grouped into four main categories, namely statistical, structural, spectral and model based feature extraction techniques (Xie and Mirmehdi, 2008). Statistical methods are based on the spatial distribution of pixel gray-level values. Structural methods are more focused on extracting the underlying texture primitives and spatial placement rules of the texture. Spectral methods are basically filtering techniques which decompose an image into sub band images. Model based methods use generative models to represent images, with the estimated model parameters as texture features.

Among the model based methods, Markov Random Field (MRF) is one of the popular generative texture feature extraction techniques. In MRF theory, a texture is assumed as a realization of a random process characterized by a specific probability model and its model parameters. MRFs have proven to offer a powerful framework for image texture analysis (Petrou and Sevilla, 2006).

Gaussian Markov Random Field (GMRF) is an important subclass of MRF whose joint probability model is a multivariate Gaussian distribution (Rue and Held, 2005). A local conditional probability model of GMRF encapsulates spatial dependencies between a pixel and its neighbours (Zhao et al., 2007). Model parameters of the conditional distribution of GMRF offer a satisfactory feature set, which successfully enables the discrimination of many different textures (Chellappa and Chatterjee, 1985; Descombes et al., 1999).

GMRF is frequently chosen for texture feature extraction over MRF because of its simplicity in parameter estimation which permits to have an analytically and computationally efficient feature extraction process (Tuceryan and Jain, 1998). Also GMRF has a well defined model form with a proper partition function and finite number of model parameters. Simplicity and efficiency are regarded as the key qualities of successful texture descriptors (Mäenpää and Pietikäinen, 2005). Effective texture segmentation and classification results have been reported using GMRF descriptors (Manjunath and Chellappa, 1991; Zhao et al., 2007; Mahmoodi and Gunn, 2011).

GMRF features describe spatial dependency between pixels which is a primary characteristic associated with texture. Even though GMRF descriptors are enriched with information about spatial interactions of the texture, they ignore the use of some important structural and statistical information about the texture such as descriptions of texture primitives and their spatial distributions. Therefore, possibility of further improvements exist which can lead to better texture descriptors. GMRF based features produce good results for homogeneous, fine, stochastic textures, but have been performed poorly when characterizing more structured and macro textures (Petrou and Sevilla, 2006).

Model based texture feature extraction methods such as GMRFs are popular in texture analysis because they are derived based on well established probabilistic models and provide generative texture features which can be used in texture synthesis. However, in texture segmentation and classification more descriptive features are favoured than the generative features unlike in a synthesis problem. Even though successful texture classification and segmentation over small texture datasets are reported in the literature, these datasets are hardly adequate to evaluate the generalized performance of GMRF features (Chellappa and Chatterjee, 1985; Manjunath and Chellappa, 1991; Zhao et al., 2007). Therefore, their generalized performance on larger texture datasets is of interest

and further techniques to improve their discriminative power for texture classification and segmentation tasks are worth giving consideration.

1.2 Scope and Contributions

The scope of this research is restricted to the GMRF model based texture feature extraction. GMRF feature extraction associates relatively low computational cost compared to other model based texture feature extraction schemes. Moreover they provide a robust way of determining spatial dependencies between a pixel and its neighbourhood which are more favourable and significant features in texture characterization.

The main contribution of this work is to introduce a novel texture descriptor based on GMRFs which overcomes certain flaws in existing traditional GMRF features. In this regard, the Local Parameter Histogram (LPH) descriptor is proposed. LPH features simply integrate both statistical and structural properties of a texture. Traditional GMRF features represent the model parameter estimates of the GMRF. The parameter estimation is achieved via Least Square Estimation (LSE). The process of estimation of GMRF model parameters results in highly over smoothed parameter estimates in respect of texture. The parameter estimates are spatially constant and thus disregard some of the important statistical and structural texture information for texture discrimination.

LPH descriptors can overcome the problem of parameter smoothing occurring in traditional GMRF parameter estimation. The proposed method suggests simple alterations to the existing GMRF feature extraction technique and achieves significantly better results. The parameter estimation is performed using the LSE similar to the traditional GMRF parameter estimation, however we fit localized linear models at each pixel based on local linear regression. The inconsistencies arising in the local parameter estimation process are addressed by applying Tikhonov regularization and introducing a rule for estimation window size selection. The localized process of parameter estimation results in spatial variations in the parameter estimates which are repetitive with the periodic texture patterns. The distributions of these local parameters are proposed as a successful discriminative texture descriptor and referred to as LPH features.

Next we introduce a rotation invariant framework for LPH descriptors to achieve invariant features. Due to the histogram construction procedure, the LPH features are clearly translation invariant and exhibit scale invariance up to a certain degree. However, they are not rotational invariant to perform invariant texture analysis. Therefore, LPH features are made rotation invariant via a local circular neighbourhood shifting procedure which leads to Rotational Invariant LPH features (RI-LPH). Another technique based on Isotropic GMRF (IGMRF) is also suggested to achieve rotation invariance which leads to Isotropic LPH features (I-LPH). I-LPH feature extraction is relatively faster and efficient, compared to RI-LPH feature extraction.

In this research, general texture segmentation and classification on large texture datasets are performed and generalized performance of proposed features with traditional GMRF features are evaluated. Commonly used Brodatz (Brodatz, 1966), Outex (Outex Texture Database, 2007), Curet (Dana et al., 1997), Prague (Haindl and Mikeš, 2008) and UIUC (Lazebnik et al., 2005) texture datasets are employed. Comparisons against the state-of-the-art structural and spectral texture feature extraction methods based on local feature distributions are also carried out, namely with uniform local binary patterns (Ojala et al., 2002) and spectral histograms (Liu and Wang, 2006). LPH, RI-LPH and I-LPH descriptors capture both pixel dependencies and their spatial distributions and thus give a significant performance improvement over traditional GMRF features. Also comparable results with the state-of-the-art methods are achieved.

Furthermore, natural image segmentation using Berkeley data (Martin et al., 2001) is performed. The simple k-means clustering algorithm is employed as the segmentation method and satisfactory segmented results are achieved. Further analysis with natural images, by integrating colour information and using advanced segmentation algorithms for better boundary localization is also examined.

Moreover, a real world application in medical image processing is considered. I-LPH features are introduced in the diagnosis and quantification of emphysema and its subtypes. Emphysema is a serious lung disease which fatally disturbs the respiratory process which is recognized as the fourth leading cause of death in the world (Sørensen et al., 2010). Results show that the novel texture features can perform well in discriminating different lung tissues, giving comparative results with the current state of the art texture based emphysema quantification. Furthermore, supervised lung parenchyma tissue segmentation is carried out and the effective pathology extents and successful tissue quantification are achieved.

Finally, a Bayesian framework to integrate prior knowledge into the parameter estimation is considered. Here a posterior model for achieving spatially varying parameters is introduced for modelling texture. Many computer vision applications exhibit some prior knowledge in addition to the visual image data. This prior information could be integrated into the texture descriptors through a Bayesian framework to achieve application specific performance improvement. Current study simply uses a smoothing prior to demonstrate the improvement using prior knowledge integrated texture descriptors.

1.3 Thesis Organization

This thesis is organized as follows. Chapter 2 gives a general review about image texture and existing texture feature extraction methods. In Chapter 3, the GMRF model, parameter estimation and its weaknesses are comprehensively discussed and current methods of GMRF feature extraction are explored. Chapter 4 introduces the novel

texture descriptor, the LPH descriptor and Chapter 5 focuses on methods of achieving rotation invariant features. Chapter 6 and Chapter 7 illustrate the texture segmentation and classification results respectively. In Chapter 8, the features are applied on to a real world medical application where emphysema tissue classification and quantification are carried out. Chapter 9 introduce the Bayesian framework for prior knowledge integration. Finally in Chapter 10 conclusions of the research are reported and plausible future work are outlined.

1.4 Publications

Publications based on this research include:

[1] Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Niranjan, M. (2012). Unsupervised texture segmentation using active contours and local distributions of Gaussian Markov random field parameters. In Proc. of British Machine Vision Conference, pages 88.1-88.11.

[2] Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Niranjan, M. (2014). Quantitative analysis of pulmonary emphysema using isotropic Gaussian Markov random fields. In Proc. of Int'l Conf. on Computer Vision Theory and Applications, pages 44-53.

[3] Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Niranjan, M. (2014). An inhomogeneous Bayesian texture model for spatially varying parameter estimation. In Proc. of Int'l Conf. on Pattern Recognition Applications and Methods, pages 139-146.

[4] Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Niranjan, M. (2014). Gaussian Markov random field based improved texture descriptor for image segmentation. Image and Vision Computing, 32: 884-895.

[5] Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Niranjan, M. Rotation invariant texture descriptors based on Gaussian Markov random fields for classification. Pattern Recognition Letters, Preprint, submitted September 17, 2014.

Chapter 2

Texture Analysis

2.1 Texture

Texture can be broadly defined as the visual or tactile surface characteristics and appearance of an object or a region. A texture exhibits a regular repetition of an element or a pattern. The patterns can be the result of physical surface properties such as roughness or oriented strands which often have a tactile quality, or they could be the result of reflectance differences such as the colour on a surface (Tuceryan and Jain, 1998). Texture contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment.

Despite its importance and extensive presence in image data, there is no unique and precise definition of texture (Patil et al., 2013; Nixon and Aguado, 2008; Tuceryan and Jain, 1998). Tuceryan and Jain (1998) have brought together a multitude of definitions for texture which has been used in the texture analysis community. Some of these definitions describe the texture as below.

- "We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule." (Tamura et al., 1978)
- "A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic." (Sklansky, 1978)
- "An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives." (Haralick, 1979)

These definitions provide the primary characteristics of a texture i.e. existence of a texture building block and the repetitiveness that forms the arrangement patterns. The

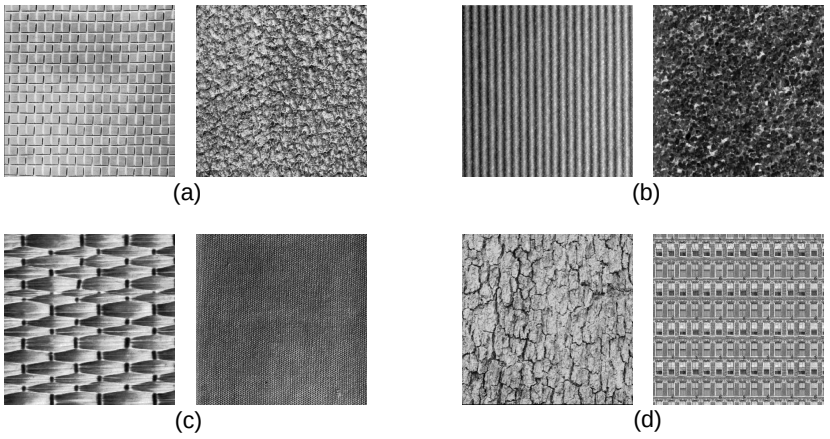


Figure 2.1: Perception of textures. (a) deterministic vs statistical, (b) directional vs non directional, (c) coarse vs fine, (d) natural vs synthetic.

repeating local structure or the the fundamental building block of the texture is known as the texture primitive or the texel (Petrou and Sevilla, 2006). The notion of repeating characteristic contributes to the homogeneity of a texture. The homogeneity implies that local characteristics of the texture do not depend on the location (Tuceryan and Jain, 1998).

In view of many definitions of texture, Karu et al. (1996) have established two common agreeable factors about a texture.

1. Texture has significant variation in intensity levels between nearby pixels. i.e., at some scale, there is non-homogeneity.

2. Texture has a homogeneous property at some spatial scale larger than the texture primitive size.

The first statement describes the local appearance of the texture in the texel level and the second statement defines the global homogeneous attribute of the texture.

Compared with intensity, texture is more of a global property where texture can only be perceived from an image region which is large enough. Depending on the variations and structural arrangements on the surfaces, textures may be perceived as being deterministic or statistical, directional or non-directional (isotropic), smooth or rough, coarse or fine, natural or synthetic, etc. Figure 2.1 illustrates these variations in textures. In deterministic textures the texel is clearly distinct, however, in stochastic textures finding an exact unique texel is difficult but the local statistical properties still show some repetitiveness. Furthermore, large primitives give rise to coarse or large textures and small primitives produce fine textures (Singh and Singh, 2002).

2.2 Image Texture Analysis

Major issues in texture analysis may be summarized as follows:

[1] ***Texture feature extraction***
Feature extraction is concerned with the quantification of texture characteristics in terms of a collection of descriptors or quantitative feature measurements, often referred to as a feature vector. The choice of appropriate descriptive features will radically influence the reliability and effectiveness of subsequent texture analysis processes (Patil et al., 2013). There exist invariance requirements for feature extraction, specifically invariance to position, scale and rotation. Clearly, position invariance is a main requirement of texture descriptors. Rotation invariance is a preferred characteristic of texture features, however it is not a strong requirement as position invariance. Scale invariance is the least strong requirement and its necessity mainly depends on the application (Nixon and Aguado, 2008).

[2] ***Texture classification***
The goal of texture classification is to design an algorithm to categorize an unknown texture image as belonging to one of a set of known classes, depending on previously known training data. The classical pattern classification techniques such as k-nearest neighbour algorithm, Bayesian classifier, support vector machines etc. are commonly employed in texture classification (Petrou and Sevilla, 2006; Nixon and Aguado, 2008; Patil et al., 2013).

[3] ***Texture segmentation***
Segmentation divides the image into different texture regions. Texture segmentation is a difficult problem because one usually does not know a priori what types of textures exist in an image, how many different textures there are, and what regions in the image have which textures. According to how much prior knowledge is involved, this problem can be divided into three subsets, namely supervised segmentation, semi-supervised segmentation, and unsupervised segmentation. In supervised segmentation, training data are available and the number of different textures in the image is known. In semi-supervised segmentation only the number of different textures in the image is known and in unsupervised segmentation non of this information is available (Xia et al., 2006b). The two general techniques of performing texture segmentation are region-based approaches or boundary-based approaches. In the region-based approaches, one tries to identify regions of the image which have a uniform texture and in the boundary-based approaches, region boundaries between different textures are detected (Tuceryan and Jain, 1998).

[4] ***Texture synthesis***
Synthesis generates or fill holes with a texture which is similar to a target texture (Efros and Leung, 1999). Approaches that synthesize texture are relatively recent compared to the history of other texture analysis approaches. Texture synthesis is a problem which

Figure 2.2: General texture analysis system.

is more popular in computer graphics where quality of the rendered scenes are improved by including texture (Nixon and Aguado, 2008). Both parametric and non-parametric methods have been developed for texture synthesis. A well-known technique to texture synthesis is to use a Markov random field (Efros and Leung, 1999; Petrou and Sevilla, 2006; Nixon and Aguado, 2008).

The basic texture analysis system in computer vision is shown in Figure 2.2. The first stage of any vision system is the image acquisition stage. If the image has not been acquired satisfactorily then the intended tasks may not be achievable, even with the aid of some form of image enhancement. Careful structuring of the lighting arrangement and camera position may be used to enhance the particular features of interest. Next the texture feature extraction process is performed which is sometimes considered as a preprocessing stage in certain applications. Finally, texture classification, segmentation or synthesis is carried out accordingly. In some occasions, before the classification stage, a feature selection or a feature smoothing process is carried out to improve the results. Moreover, if necessary, an additional pre-processing stage may be conducted prior to the feature extraction to enhance the quality of the input image.

In this research, our primary focus is on texture feature extraction for texture classification and segmentation purposes. Texture classification and segmentation are fundamental problems in computer vision with a variety of applications (Liu and Wang, 2003; Tou et al., 2009; Zhang et al., 2002). Descriptive texture features are extensively employed in many applications, for example in image retrieval, remote sensing, medical image processing, object recognition, surface inspection and document processing. Unlike in a synthesis problem, for texture classification and segmentation, discriminative features are required rather than reliable modelling of the texture (Zhao et al., 2007).

2.3 Current Methods of Texture Feature Extraction

Texture analysis has a rich history in image processing and computer vision (Nixon and Aguado, 2008). A large number of texture feature extraction methods have been developed. These methods have been grouped into different categories in numerous ways according to some of the basic characteristics. One such categorization is as follows.

[1] ***Statistical methods***
Statistical methods are based on the spatial distribution of pixel gray-level values. They

are well rooted in the computer vision literature and have been extensively applied to various tasks. Texture features are computed based on the statistical distribution of image intensities at specified relative pixel positions. A large number of these features have been proposed, ranging from first order statistics to higher order statistics.

[2] ***Structural methods***
Structural methods are more focused on extracting the underlying texture primitives and spatial placement rules of the texture. From the structural point of view, texture is characterized by the texels or texture primitives, and the spatial arrangement of these primitives. The texture primitive can be as simple as individual pixels, a region with uniform gray-levels, or line segments. The placement rules can be obtained through modelling geometric relationships between primitives or learning their statistical properties.

[3] ***Spectral methods***
Spectral methods are basically filtering techniques which decompose an image into sub band images. The texture features can be derived from the spatial domain, the frequency domain and the joint spatial/spatial-frequency domain. In the spatial domain, the images are usually filtered by gradient filters to extract edges, lines, isolated dots, etc. Many other features related to frequency spectrum have been derived by filter response image. It has been found that the human visual system transforms the retinal image into sequence of sub band representations which can be mathematically modelled by convolving the input image with a bank of filters with tuned frequencies and orientations. Following this notion, spectral features are commonly extracted by applying filter banks to decompose an image (Movellan, 2002).

[4] ***Model based methods***
Model based methods generally use stochastic and generative models to represent images, with the estimated model parameters as texture features. This category of methods are predominantly used in extracting generative features which can be applied in texture synthesis. However, later on it has been also applied in texture classification and segmentation tasks (Petrou and Sevilla, 2006).

A large number of texture feature extraction methods have been discussed in broad surveys by Tuceryan and Jain (1998); Zhang and Tan (2002); Petrou and Sevilla (2006); Xie and Mirmehdi (2008); Srinivasan and Shobha (2008). Some of the popular texture feature extraction techniques are reviewed here under the above categorization which have been used in texture classification and segmentation. Following subsections briefly review some popular methods providing comprehensive details on the state-of-the-art methods. Comparisons of the methods are discussed in a separate subsequent section.

2.3.1 Statistical Methods

The most simple and efficient statistical texture feature extraction technique is the image intensity histogram. The image intensity histogram is a first order statistical feature that is not only computationally simple, but also rotation and translation invariant. First order texture measures are statistics calculated based on individual pixel intensity values and do not consider pixel neighbourhood relationships. Common features extracted from intensity histograms include moments such as mean, variance, dispersion, mean square value or average energy, entropy, skewness and kurtosis (Srinivasan and Shobha, 2008). Despite their simplicity, histogram techniques have proved their worth as a low cost, low level approach (Swain and Ballard, 1992; Hadjidemetriou et al., 2003). Nevertheless, texture analysis based solely on the gray-level histogram suffers from the limitation that it provides no information about the spatial interactions and arrangements between pixels. Therefore, higher order feature statistics have been introduced.

Autocorrelation features are a second order feature statistic which measures the correlation between the texture image itself and a translated version of the image with a given displacement vector. The autocorrelation function can be used to assess the amount of regularity as well as the fineness/coarseness of the texture present in the image. However, autocorrelation statistic is sensitive to noise interference. The higher order correlations relatively perform better than lower order correlations (Coroyer et al., 1997; Huang and Chan, 2004; Petrou and Sevilla, 2006).

Spatial Gray-Level Co-occurrence Matrices (GLCM) are another well known texture feature introduced by Haralick et al. (1973). GLCM estimates image properties related to second-order statistics. These second order statistics are accumulated into a set of 2D matrices, each of which measures the frequency of occurrence of two gray-levels separated by a given distance. Haralick et al. (1973) has proposed a number of useful texture features that can be computed from the GLCM such as energy, entropy, contrast and correlation. Some shortcomings of GLCM is that when constructing GLCM, the number of gray-levels should be reduced in order to keep the size of the GLCM manageable. It is important to ensure the number of entries of each matrix is adequate to be statistically reliable. For a given displacement vector, a large number of features can be computed, therefore feature selection procedures are required (Petrou and Sevilla, 2006).

The gray-level run length is introduced by Galloway (1975). A run is defined as consecutive pixels with the same gray-level, collinear in the same direction. The number of pixels in a run is referred to as run length, and the frequency at which such a run occurs is known as run length value. Some of the statistics commonly extracted from run length matrices for texture analysis are short run emphasis, long run emphasis, etc. (Xie and Mirmehdi, 2008).

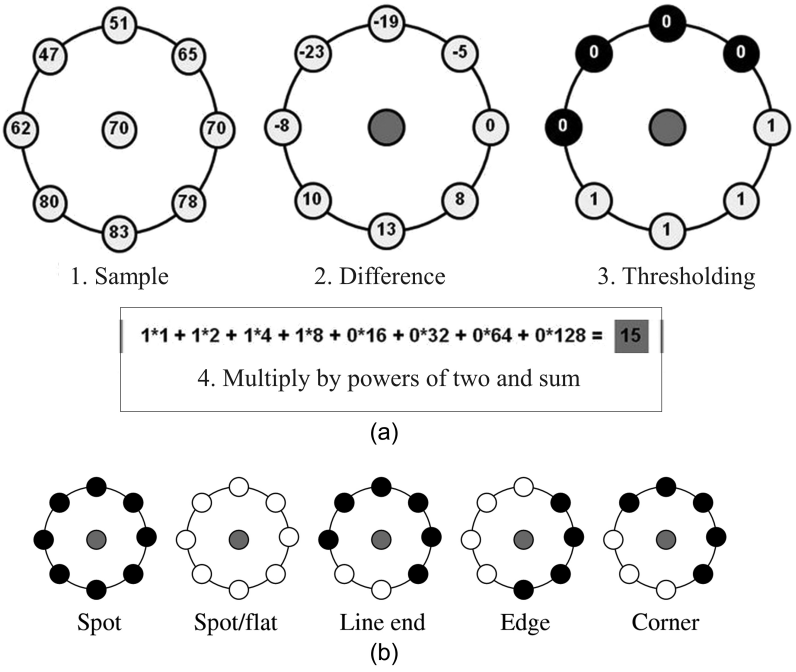


Figure 2.3: Local binary patterns. (a) basic construction stages (Pietikäinen, 2010), (b) different texture primitives detected by LBP (Mäenpää and Pietikäinen, 2005).

2.3.2 Structural Methods

The local binary pattern (LBP) is the currently most popular structural feature extraction technique. LBP texture operator is first introduced as a complementary measure for local image contrast (Ojala et al., 1996). Later, a more general formulation was proposed that further allowed for multi-resolution and rotation invariant analysis (Ojala et al., 2002). The LBPs are obtained by thresholding neighbour pixel values $G_P = \{g_p | p = 0, \ldots, P-1\}$ in a local neighbourhood with respect to the intensity g_c of the centre pixel at (x_c, y_c) and is given by,

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s\left(g_p - g_c\right) 2^p \tag{2.1}$$

where P is the number of neighbours in a circular neighbourhood situated at a radius R from (x_c, y_c), and $s(.)$ is the Heaviside function. LBP are invariant to any monotonic gray-scale transformation. Fig 2.3a illustrates the basic stages of LBP feature extraction. Note that, by choosing a fixed reference position on the neighbourhood as the "leading bit" the thresholded neighbour values can be interpreted as bits and a P bit binary number can be computed. The LBP measures the local structure by assigning unique identifiers, the binary number, to various micro-structures or texture primitives in the image. Different LBP codes assigned to texture primitives are shown in Figure 2.3b.

By varying the radius and the number of neighbours, the structures are measured at different scales. Rotation invariant LBP are achieved by rotating the circular neighbourhood until the lowest possible binary number is found.

$$LBP^{ri}_{P,R} = \min\{ROR(LBP_{P,R}, i) | i = 0, 1, \ldots, P-1\} \tag{2.2}$$

where $ROR(b, i)$ performs i circular bit-wise right shifts on P-bit binary number b. It is observed that certain patterns seem to be fundamental properties of texture, providing the vast majority of patterns, sometimes over 90% (Mäenpää and Pietikäinen, 2005). These patterns are called uniform because they have one thing in common, i.e. at most two one-to-zero or zero-to-one transitions in the circular binary code. The LBP codes shown in Figure 2.3b are all uniform. The uniformity of a binary number x is defined by,

$$U(x) = \sum_{p=0}^{P-1} F\left(x \;\; XOR \;\; ROR(x, 1), p\right) \tag{2.3}$$

where the function $F(x, i)$ extracts the i^{th} bit from a binary number x. Therefore, the uniform local binary patterns are defined as,

$$LBP^{ri2}_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) & U(G_P) \leq 2 \\ P + 1 & \text{otherwise} \end{cases}$$

LBP is a simple yet very efficient texture operator with computational simplicity. LBP descriptors have been widely used in many fields including texture analysis and is currently one of the state-of-the-art methods in image feature extraction (Topi et al., 2000; Liao and Chung, 2007; Ahonen and Pietikäinen, 2008; Sørensen et al., 2010; Guo et al., 2010; Guo and Zhang, 2010; Waller et al., 2011, 2012).

One of the other frequently used structural texture feature extraction method is texton features. Textons were first introduced by Julesz (1981) as the fundamental image structures and they are considered as atoms of pre-attentive (unconscious) human visual perception. Leung and Malik (2001) adopted a discriminative model to describe textons. Each texture image is analysed using a filter bank composed of 48 Gaussian filters with different orientations, scales and phases. Thus, a high dimensional feature vector is extracted at each pixel position. K-means is used to cluster those filter response vectors into a few mean vectors which are referred to as textons. Varma and Zisserman (2002) also adapted a similar technique to extract textons, where instead of filter responses they have used image patches which are closely described by a non-parametric Markov random field. Zhu et al. (2005) also present a generative image model for learning textons from texture images. Blostein and Ahuja (1989) examine the response of the Laplacian of Gaussian (LoG) filter at multiple scales to extract texture primitives.

2.3.3 Spectral Methods

Different kind of filters have been used separately to filter images and extract features. Some popular filters are Laws operators, difference of Gaussians filters, derivative of Gaussian filters, Laplacian of Gaussian filters, steerable filters, Gabor filters (Laws, 1980; Tuceryan and Jain, 1998; Leung and Malik, 2001; Tou et al., 2009; Lowe, 2004; Varma and Zisserman, 2009; Griffin and Lillholm, 2010). Specially designed multi-channel filter banks are a popular technique used for texture feature extraction (Jain and Farrokhnia, 1990; Bosnjak et al., 1998; Zhang et al., 2002; Varma and Zisserman, 2002; Recio R. et al., 2005; Lee et al., 2006; Sagiv et al., 2006; Bianconi and Fernandez, 2007; Tou et al., 2007; Olowoyeye et al., 2009; Veni, 2010).

Pyramid based methods are employed to extract texture features over multiple scales. The pyramid of filter responses based on a certain type of selected filters is employed to achieve this task. The n^{th} level image I^n of the pyramid is usually constructed by down-sampling (or up-sampling) the filter response of the image in the previous level I^{n+1}. Features derived based on the filter pyramid have been used as multi-scale texture features. The Gaussian pyramid is one of the simplest multi-scale transforms. The finest scale layer (first level) is the original image. As each level is a low pass filter version of the previous level, the low frequency information is repeatedly represented in the Gaussian pyramid (Xie and Mirmehdi, 2008). The Laplacian pyramid introduced by Burt and Adelson (1983) which use Laplacian of Gaussian filter bank instead of Gaussian filters. The oriented pyramid is another multi-scale technique which decomposes an image into several scales and different orientations. Unlike the Laplacian pyramid where there is no orientation information in each scale, in an oriented pyramid each scale represents textural energy at a particular direction. One way of generating an orientated pyramid is by applying derivative filters to a Gaussian pyramid or directional filters to a Laplacian pyramid (Xie and Mirmehdi, 2008). A steerable pyramid is another way of analysing texture in multiple scales and different orientations (Simoncelli and Freeman, 1995).

The power spectrum is another spectral feature extraction technique. The power spectrum represents the energy distribution in the frequency domain. It is commonly generated using the discrete Fourier transformation. The radial distribution of energy in the power spectrum reflects the coarseness of the texture, and the angular distribution relates to the directionality. Thus, the power spectrum can be used to characterise textures. Commonly used feature extraction techniques include applying ring filters, wedge filters, and peak extraction algorithms on the power spectrum (Gonzalez et al., 2004).

Spectral histogram is another popular and efficient choice in spectral feature extraction. A spectral histogram is a feature vector consisting of the marginal distribution of filter responses. Given a texture image I and a bank of filters $\{\mathbf{F}^{(\alpha)}, \alpha = 1, 2, \ldots, K\}$, for each filter $\mathbf{F}^{(\alpha)}$ a sub band image $I^{(\alpha)}$ through linear convolution is computed. i.e. at a

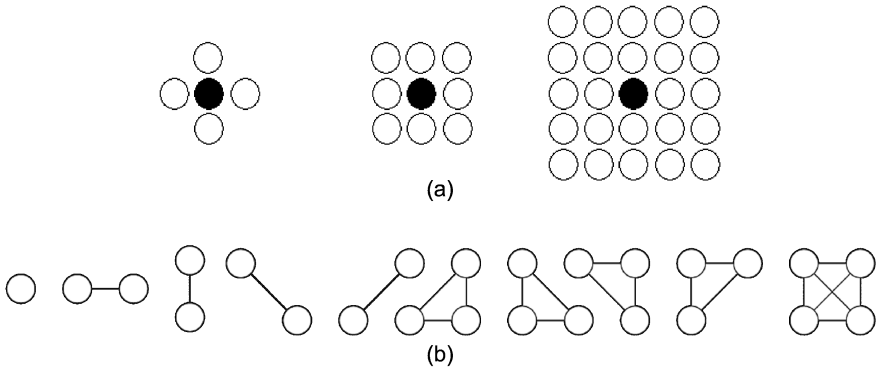


Figure 2.4: (a) Different orders of neighbourhoods, from left first order, second order and eighth order neighbourhoods, (b) all required clique types for MRF models up to the eighth order (Blunsden, 2004).

pixel location v,

$$I^{(\alpha)}(v) = \mathbf{F}^{(\alpha)} * I(v) = \sum_{u} F^{(\alpha)}(u) I(v-u) \tag{2.5}$$

The histogram of $I^{(\alpha)}$ is defined as,

$$H_I^{(\alpha)}(z) = \frac{1}{|D|} \sum_{v} \delta(z - I^{(\alpha)}(v))$$

Where $\delta()$ is dirac delta function. The spectral histogram with respect to the given filters is then given by,

$$H_I = (H_I^{(1)},\ H_I^{(2)}, \ldots,\ H_I^{(K)}) \tag{2.6}$$

The spectral histogram is translation invariant which is often a desirable property in texture analysis and with a sufficient number of filters it can uniquely represent any image up to a translation (Liu and Wang, 2003). The spectral methods capture macro level non uniform information relatively better than structural methods.

2.3.4 Model based methods

Markov random fields (MRF) are one of the key model based texture analysis techniques. They have been used as a tool for image and texture synthesis. Features extracted from MRF are both descriptive and generative. Thus they have been found to be useful in texture classification, segmentation and other computer vision applications (Held et al., 1997; Wang, 1999; Huang et al., 2003; Xia et al., 2006b; Gomez and Salinas, 2006; Li, 2009). A texture can be considered to be generated by a certain random process. MRF is a probability model which provides a convenient way to model local spatial interactions among entities such as pixels. In MRF texture modelling to model a texture is to specify the corresponding conditional probabilities and estimating their parameters.

In MRF models, an image is represented by a finite rectangular lattice S within which each pixel is considered as a site $s \in S$. From the Markov property it is assumed that the intensity of a pixel is conditionally dependent upon the local neighbourhood N and independent of the rest of the image. Figure 2.4a illustrates some example neighbourhoods of MRFs. An MRF can be characterized by a Gibbs distribution. The Gibbs distribution is usually defined with respect to *cliques*. A clique is defined as a subset C in an image lattice S where every pair of distinct pixels in C are neighbours of each other. A set of cliques $\mathbf{C}$ is shown in Figure 2.4b. The Gibbs distribution with respect to a given neighbourhood N is given by,

$$p(I) = \frac{1}{Z} \exp\{-E(I)\} \tag{2.7}$$

where Z is the normalization factor called partition function and $E(I)$ is the energy function. The $E(I)$ is given by sum of clique potentials $V_C(s)$ as,

$$E(I) = \sum_{C \in \mathbf{C}} V_C(s) \tag{2.8}$$

$V_C(s)$ is a function of pixels belong to the clique C. Usually single site and pair site clique potentials are used in texture modelling (Tuceryan and Jain, 1998; Li, 2009; Petrou and Sevilla, 2006). $p(I)$ defines the joint probability model of MRF. Because of the assumption of local dependency, the joint probability model for a random field could be factorized into clique potentials. This reduces the complexity of a global image modelling problem and lead to the local probability model. One can model the texture either globally by specifying the total energy of the lattice or model it locally by specifying the local interactions of the neighbouring pixels in terms of the conditional probabilities. Depending on how the energy function is selected, there exist different types of MRFs, for example, auto-models, multi level logistic model etc. A comprehensive discussion of different types of MRFs can be found in Li (2009) and we will explain GMRFs in detail in the next chapter.

Although the parametric MRF texture analysis has a well defined mathematical model, parameter estimation is a computational expensive process. Also the partition function of MRFs is usually difficult to find and approximations are needed. The GMRF and isotropic Gaussian MRF have been suggested which are powerful tools of texture modelling which involve significantly less computational expense (Chellappa and Chatterjee, 1985; Kashyap and Khotanzad, 1986).

A non parametric MRF model introduced by (Varma and Zisserman, 2009) is another approach of overcoming the computational difficulties while employing effective MRF concepts. This model associates a supervised scheme where a texton dictionary is first learnt which is then used to develop a conditional distribution to describe the texture. Other non parametric MRF model based approaches have been also introduced and

efficiently used in texture analysis (Efros and Leung, 1999; Blunsden, 2004; Kwatra et al., 2005; Blunsden and Torrealba, 2005).

2.3.5 Comparisons

There have been many studies comparing various subsets of texture features. The results in most of these work much depend on the data set used, the set of parameters used for the methods examined, and the application domain (Xie and Mirmehdi, 2008).

Ohanian and Dubes (1992) compared the fractal model, GLCM, the MRF model and Gabor filtering for texture classification. The GLCM features generally outperformed other features in terms of classification rate. Singh and Singh (2002) have also shown that GLCM perform better than other statistical texture features. GLCM has also performed better than some structural texture features, given that higher number of features derived from GLCM are employed (Patil et al., 2013).

In another related work, Pichler et al. (1996) reported superior results using Gabor filtering over other wavelet transforms such as pyramidal and tree structured wavelet transforms. Spectral methods are very popularly used in the recent years, especially the Gabor filters. Although this method requires selection of a proper filter set, the accuracy achieved is good and usually outperform simple statistical techniques (Zhang and Tan, 2002; Patil et al., 2013). Spectral histograms and pyramid based methods have been successfully employed in texture classification, segmentation and also in texture synthesis (Heeger and Bergen, 1995; Zhu et al., 1998; Liu and Wang, 2002, 2003; Long and Younan, 2006; Liu and Wang, 2006).

Structural methods specially LBP method has been shown superior texture classification performances compared to model based GMRF features and GLCM features (Ojala et al., 2001, 2002). Varma and Zisserman (2002) demonstrated superior performance of texton clustering and distribution features constructed using Gaussian, Laplacian of Gaussian and oriented filter responses. However, compared to texton, construction of LBP is efficient and does not require any training data in the feature construction process. The LBP method has been widely used in texture analysis because of its excellent property of gray-scale invariance and the high discriminative power despite of the theoretical simplicity (Mäenpää and Pietikäinen, 2005).

In general, statistical approaches are generally better suited to micro textures, and structural methods are better perform on textures that exhibit a regular macro-structure. Spectral methods generally perform well for all types of textures, however, there is no unique way of selecting a optimal filter set. Model based approaches support different

kinds of textures depending on the nature of the model. In conclusion, spectral histograms and LBP features have been widely employed in texture analysis with satisfactory performances and can be considered as the current state-of-the-art texture features. It is interesting to note that these methods are based on local feature distributions.

Although model based features such as MRF, GMRF, isotropic Gaussian MRF have been reported to produce successful results on small scale selected texture datasets, in general, model based techniques have shown lower performance compared to other non-model based techniques specially in texture classification and segmentation (Ojala et al., 2001; Singh and Singh, 2002; Xie and Mirmehdi, 2008). Also model based techniques associate relatively higher computational costs.

2.4 Conclusions

Texture feature extraction is an important task in image processing and computer vision. Among numerous texture feature extraction methods introduced in the literature, LBP and spectral histogram features have shown a superior texture discriminative power and have been extensively used in texture analysis. In general, statistical, structural and spectral methods provide more efficient descriptive features relative to the model based generative texture features. These non-model based techniques are very popular in texture classification and segmentation and have also been adapted for use in texture synthesis in recent years. However, model based texture features such as GMRF features are less popular due to their limited discriminative power and later developed more efficient discriminative non-model based features. Comparisons between GMRF and other non-model based features have been carried out and due to its poor performance GMRF features are hardly used in the present except in a few application specific tasks. This motivates the current research to analyse weaknesses in GMRF features and to develop their potential for better texture segmentation and classification. GMRFs are mathematically well established modelling tools in probability and statistical theories which can model contextual dependencies in an image. Local dependencies could be one of the key properties for texture recognition and GMRF provides the statistical means for quantifying local interdependencies. However, GMRF features lack the ability to capture some of the important structural and statistical features in its current method of formulation. Therefore, GMRF features could be further improved to achieve better, meaningful and competitive texture features.

Chapter 3

Gaussian Markov Random Fields

3.1 Random Fields and Markov Random Fields

A two dimensional (2D) random field is a collection of random variables which are arranged in 2D space. An image for example, is a random field on a 2D regular lattice Ω. Each lattice point or a pixel location is generally referred to as a site $s = \{i, j\}$, $s \in \Omega$. If a site s is associated with a random variable y_s for example, intensity value, then the random field Y is given by,

$$Y = \{y_s | s \in \Omega\} \tag{3.1}$$

A random field, in other words, also represents a stochastic process where instead of in time the random variables are arranged in space. For the sake of probabilistic interpretation, Y is represented by a multidimensional random vector $Y = [y_{s_1}, y_{s_2}, \ldots, y_{s_K}]^t$ where K number of random variables signify the lattice (Bouman, 2009). Therefore, the joint probability model of the random field can be represented as $p(Y)$ and the full conditional probability model can be written as $p(y_s | y_t, t \neq s, t \in \Omega)$.

Markov Random Field (MRF) is a type of random field which exhibits the Markovian property. The Markovian property states that the value of a pixel is conditionally independent from all the other pixels of the random field except for its neighbouring pixels (Petrou and Sevilla, 2006). Using the Markovian property the full conditional probability model of the random field can be narrowed down to the following,

$$p(y_s | y_t, t \neq s, t \in \Omega) = p(y_s | y_{s+r}, r \in N) \tag{3.2}$$

where the set N defines the relative coordinates of the neighbourhood set with respect to the site s. The equation (3.2) implies the dependency of a random variable y_s on its neighbours. From the point of view of a texture, because of the repetitive nature,

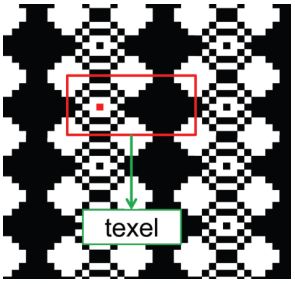


Figure 3.1: The texel of a deterministic texture.

a texture has the dependency on the local characteristics. For example, considering a deterministic texture, a pixel value in the texture will only depend on its neighbours who are located in the texel (See Figure 3.1). Thus, the MRF model is greatly appropriate for texture modelling and has been widely used in stochastic modelling of textures (Kim et al., 2006; Cesmeli and Wang, 2001).

MRFs provide a convenient way of statistical modelling of texture, globally by means of joint probability distributions and locally by conditional probability distributions. A particular MRF model means a particular probability function specified by a particular functional form and its parameters (Li, 2009). The joint probability distribution of the MRF is implicitly fully define by its conditional probability distributions (Rue and Held, 2005). Therefore, the local conditional model of the MRF model has become more popular in texture feature extraction, compared to the global joint probability model because of the computational efficiency and relative mathematical simplicity (Manjunath and Chellappa, 1991; Zhao et al., 2007). The parameters can be locally estimated and can be computed in a massively parallel manner (Li, 2009).

3.2 Conditional Gaussian Markov Random Fields

Gaussian Markov Random Field (GMRF) is an important subclass of MRFs having a finite dimensional random vector Y with the joint distribution $p(Y)$ which is a multivariate Gaussian probability density function (Rue and Held, 2005). The conditional probability distribution of GMRF model, which also takes the Gaussian form, encapsulates the spatial dependencies between a pixel and its neighbours (Zhao et al., 2007). The model parameters of the conditional GMRF is widely employed in characterizing image textures (Descombes et al., 1999). In this study we are basically focused on the conditional GMRF model also known as the local model.

Let $\Omega = \{s = (i, j) | 1 \leq i \leq H, 1 \leq j \leq W\}$ represents the set of grid points on a $H \times W$ regular lattice corresponding to an image region. The image region on Ω is pre-processed to have zero mean. The intensity value of the pixel at the location s is given by y_s and N denotes the set of relative positions of its neighbours. Then the local conditional

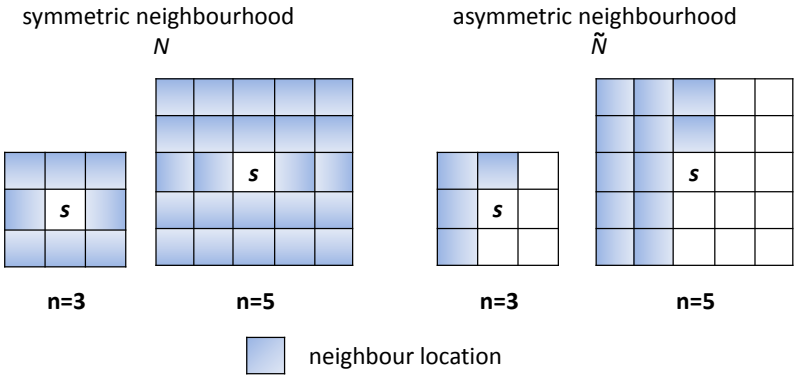


Figure 3.2: Square neighbourhoods of $n = 3$ and $n = 5$, and their corresponding asymmetric neighbourhoods

probability density function has the form,

$$p(y_s|y_{s+r}, r \in N) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y_s - \sum_{r\in N} \alpha_r y_{s+r}\right)^2\right\} \tag{3.3}$$

where α_r is the interaction coefficient which measures the influence on a pixel by a neighbour intensity value at the relative neighbour position r. For simplicity only the square neighbourhoods of size $n \times n$ pixels are used in this study for N and n is a positive odd integer value. The neighbourhood size, n is referred to as the model order.

The pixels in symmetric positions about pixel s are assumed to have identical parameters (Petrou and Sevilla, 2006; Bouman, 2009). i.e. $\alpha_r = \alpha_{-r}$ with $r \in \tilde{N}$ where $\tilde{N}$ is the asymmetric neighbourhood such that if $r \in \tilde{N}$, then $-r \notin \tilde{N}$ and $N = \{r|r \in \tilde{N}\} \cup \{r| - r \in \tilde{N}\}$ (Zhao et al., 2007). The square neighbourhoods of $n = 3$ and $n = 5$, and their corresponding asymmetric neighbourhood set are shown in Figure 3.2. Therefore, the number of interaction parameters in the model is halved and is equal to $(n^2 - 1)/2$.

The reduced model is given by,

$$p(y_s|y_{s+r}, r \in N) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y_s - \sum_{r\in \tilde{N}} \alpha_r \bar{y}_{s+r}\right)^2\right\} \tag{3.4}$$

where $\bar{y}_{s+r} = (y_{s+r} + y_{s-r})$.

For a model based approach to be successful, specially in texture classification and segmentation domains, there must exist a reasonably efficient and appropriate parameter estimation scheme, and the model itself should be parsimonious, i.e. use small number

of parameters. The conditional GMRF model employs a finite number of parameters to characterize a texture (Zhao et al., 2007). The advantage of employing GMRF model over MRF model is that the GMRF model is more simple and provides simplified ways of parameter estimation with less computational burden compared to that of MRF model (Rue and Held, 2005; Manjunath and Chellappa, 1991).

3.3 Parameter Estimation

The model parameters of the conditional GMRF model are estimated using least squares estimation (LSE) or maximum likelihood estimation (MLE) (Mahmoodi and Gunn, 2011; Manjunath and Chellappa, 1991). Note that for a GMRF model, it can be shown that MLE and LSE lead to the same set of equations (Myung, 2003; Petrou and Sevilla, 2006).

The choice of using LSE for parameter estimation was suggested by Manjunath and Chellappa (1991) which has been followed by many subsequent researchers in the field due to its fast computation properties (Stan et al., 2002). Therefore, here the LSE technique is employed for the parameter estimation. The main assumption behind the LSE method is that because the probability distribution in equation (3.4) is Gaussian, the estimated value of y_s, $\hat{y}_s$ is more probable to be the mean value of the function (Petrou and Sevilla, 2006). Therefore, the error or the residual will be,

$$\epsilon_s = y_s - \hat{y}_s = y_s - \sum_{r \in \tilde{N}} \alpha_r \bar{y}_{s+r} \tag{3.5}$$

For least squares fitting, given a stationary texture, sample neighbourhoods of the texture are extracted by linear scanning of the region Ω. Overlapping neighbourhoods are also allowed (Li, 2009). Let the interaction parameter vector be $\boldsymbol{\alpha} = col[\alpha_r | r \in \tilde{N}]$ and the neighbour value vector at the location s be $\bar{\boldsymbol{y}}_s = col[\bar{y}_{s+r} | r \in \tilde{N}]$. *col* stands for a column of elements. Then the least square solution is,

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \sum_{s \in \Omega} \epsilon_s^2 = \arg\min_{\boldsymbol{\alpha}} \sum_{s \in \Omega} \left(y_s - \boldsymbol{\alpha}^T \bar{\boldsymbol{y}}_s \right)^2 \tag{3.6}$$

The minimization problem in equation (3.6) is a convex optimization problem where analytical solutions can be easily deduced. Analytical solutions are often more efficient than equivalent numeric implementations and this is a major advantage of using GMRFs. By setting the first derivative of the residual sum of squares to zero the parameter values can be obtained as,

$$\hat{\boldsymbol{\alpha}} = \left[\sum_{s \in \Omega} \bar{\boldsymbol{y}}_s \bar{\boldsymbol{y}}_s^T \right]^{-1} \left[\sum_{s \in \Omega} \bar{\boldsymbol{y}}_s y_s \right] \tag{3.7}$$

The variance parameter of the model is then calculated by,

$$\hat{\sigma}^2 = \frac{1}{|\Omega|} \sum_{s \in \Omega} \left(y_s - \hat{\boldsymbol{\alpha}}^T \bar{\boldsymbol{y}}_s \right)^2 \tag{3.8}$$

where $|\Omega|$ is the number of sample neighbourhoods used in the estimation process, or in other words number of overlapping $n \times n$ regions extracted from the image region Ω, called the estimation window. The model parameters $\hat{\boldsymbol{\alpha}}$ and $\hat{\sigma}$ are constant over the domain Ω for a particular stationary texture. The model parameter vector $\mathbf{f} = [\hat{\boldsymbol{\alpha}}^T, \hat{\sigma}]^T$ characterizes the texture inside the region of Ω. The interaction parameters $\hat{\boldsymbol{\alpha}}$, signify the spatial inter dependencies of a pixel with its neighbours and the variance parameter $\hat{\sigma}$ indicates the roughness of the texture.

The model parameters of the conditional GMRF, i.e. interaction parameters and the variance parameter, have been employed as an effective texture feature set in texture segmentation and classification tasks. These features are called the classical or traditional GMRF features (TGMRF) (Zhao et al., 2007; Cesmeli and Wang, 2001).

In the texture classification problem, some instances of texture images are available. A texture feature vector for each given texture image is extracted separately using the above parameter estimation process. The Ω region will comprise the entire image instance. The extracted feature vectors are used to develop the training and testing sets and classification is performed subsequently.

In a texture segmentation problem, however, an image comprising different texture regions is available. According to how much a priori knowledge is involved, this problem can be divided into three subtypes, supervised, semi-supervised, and unsupervised segmentation. Feature-based segmentation algorithms can be briefly regarded as consisting of two successive processes, feature extraction and feature clustering. For each pixel, a feature vector is generated based on the local texture content over a window centred on that pixel which is called the estimation window (Xia et al., 2006b; Li, 2009). In this research a square estimation window of size w is employed. Thus, the Ω region discussed in the estimation process will become the area inside the estimation window of size $w \times w$ pixels. The parameter estimation process is then carried out using the samples extracted from the estimation window. These TGMRF features, $\mathbf{f_s}$ are referred to as adaptive TGMRF descriptors (Xia et al., 2006b; Zhao et al., 2007).

The texture inside the estimation window should be homogeneous and properly represent the complete texture pattern. Thus, larger texture patterns require larger estimation windows. Furthermore, to obtain consistent estimates the number of samples should be sufficiently greater than the number of model parameters (Petrou and Sevilla, 2006). This means the estimation window size should be large enough to provide adequate number of neighbourhood samples. Hence, using larger estimation window sizes is more common in extracting adaptive TGMRF feature sets (Zhao et al., 2007; Descombes

et al., 1999). Using large estimation windows in construction of TGMRF features has its own disadvantages which is one of the key area of consideration in the current research which will be comprehensively examined in later sections of this chapter and in the next chapter.

3.4 GMRF Texture Feature Extraction

In this section, existing GMRF feature extraction methods are briefly reviewed, specifically focusing on texture segmentation and classification. Also techniques suggested to improve the TGMRF features are also considered. Studies on this direction are rare and the datasets used in these studies are relatively small or application specific. Nevertheless, exciting developments have been proposed despite the increase of computational cost.

3.4.1 TGMRF Descriptors

In 1985 Chellappa and Chatterjee (1985) introduced the conditional GMRF features for texture classification. The least square estimates of model parameters have been used as texture features achieved from a fourth order GMRF. The classification experiment on a texture dataset based on Brodatz textures (Brodatz, 1966), which includes seven different texture classes with 16 samples for each class, reported 99.1% classification accuracy. Comparison with sample correlation features, which gives 93% accuracy for the considered dataset, demostrates the better performance of GMRF features. Minimum distance classifier has been used. Convincing texture synthesis have also been performed based on statistical fine textures.

Manjunath and Chellappa (1991) extract adaptive GMRF descriptors for unsupervised texture segmentation. Usually segmentation of textured scenes into different classes is a challenging task in computer vision where prior knowledge is unavailable. Manjunath and Chellappa (1991) have used model parameters of second order GMRF model estimated using LSE on non overlapping sub images of size 32×32 as texture features. After extracting the texture features, a two stage segmentation algorithm which performs a crude segmentation via k-means clustering algorithm and further applies a fine segmentation process by a relaxation method has been employed. They report convincing segmentation results and show that second order GMRF features are a good feature representation of real textures like wood, wool, water etc. which follow the Gaussian notion.

Xia et al. (2006b) have performed semi supervised texture segmentation where the number of texture patterns is known but the information about their properties is absent.

Adaptive second order GMRF features are employed which are constructed using an estimation window of size 21×21. They have proposed an iterative segmentation scheme which simultaneously performed parameter re-estimation and segmentation. For segmentation 12 different Brodatz textures are employed (Brodatz, 1966). By combining these textures with each other, 66 two-region mosaic images and 495 four-region mosaic images are constructed for the segmentation experiments. The segmentation error rate of 2.38% and 5.84% have been achived for the two-region and four-region mosaic image datasets respectively. They have also demostrated that the technique can be also used in successful natural image segmentation.

Mahmoodi and Gunn (2011) introduce GMRF features in an active contour based segmentation framework. The second order GMRF parameters estimated using MLE are employed. In this method simultaneous parameter estimation and segmentation is performed. Mosaic images comprising Brodatz textures (Brodatz, 1966) have been used in segmentation experiments. Convincing unsupervised segmentation results with an average segmentation error of 1.51% are achieved. They have shown that second order GMRF features from a single resolution scheme are adequate for successful segmentation via active contours algorithm.

The conditional GMRF model parameter estimates have been also used as successful texture features in real world application such as urban area extraction from Synthetic Aperture Radar (SAR) images and online defect detection systems (Xia et al., 2006a; Huang et al., 2011).

Based on the existing literature, the model parameter estimates of conditional GMRF model, which we referred to as TGMRF features, provides a discriminative texture feature in classification and segmentation. However, large scale texture evaluations are rarely performed including many different types of textures. Large estimation window sizes are involved in the estimation process and lower order GMRFs are more common due to the computational expenses. Also advanced segmentation techniques are adapted to achieve better boundary localization in segmentation tasks.

3.4.2 Improved GMRF Descriptors

Improved GMRF feature extraction methods aim to address certain weaknesses in TGMRF features. These techniques attempt to enhance the discriminative power of TGMRF features either by modifying the parameter estimation process, or constructing features derived based on TGMRF features which may achieve rotation and scale invariant features.

- Priority Sequence GMRF descriptors

The TGMRF descriptors assume that all of the neighbouring pixels, which are treated equally, interact on the centre pixel simultaneously. However it is rational to think that

neighbours have influence on center pixel in a priority sequence. i.e. closer neighbours have a higher priority to influence the centre pixel. Following this idea Zhao et al. (2007) suggest a step by step least square method to estimate model parameters. Parameters are estimated through level by level fashion from lower levels to higher levels where a level is composed of a group of pixels located similar distances away from centre pixel. Parameters in lower level neighbourhoods (near to center pixel) are computed independent of those in the higher levels and parameters of higher levels are computed based on parameters of lower level. Adaptive texture segmentation on high resolution remote sensing data is carried out to segment six land-cover classes including crop areas and residential areas. The fourth order conditional GMRF model is employed. An estimation window of size 17×17 is used to trade-off between achieving better boundary localization and achieving a stationary parameter estimation process. The overall accuracy achieved using TGMRF features is reported as 41% while the proposed features give 87.3% accuracy. This is a significant improvement over TGMRF features. However, feature selection is required to select the optimal subset of features and an iterative relaxation algorithm has to be employed in parameter estimation which leads to relatively higher computational burden.

- Hierarchical formulation

In order to improve discriminative power of TGMRF descriptors Kim et al. (2006) considered a hierarchical formulation of the features. The model parameters estimated from conditional GMRF model inherits a certain degree of fluctuation for a given texture depending on the estimation window size. This will negatively affect the subsequent segmentations. Kim et al. (2006) consider interaction coefficients of GMRF model as another random field called Random Spatial Interaction (RSI) model and use its parameter estimates as features. They have stated that parameter estimates on RSI model have less fluctuations independent of the choice of estimation window size. The segmentation results on Brodatz textures (Brodatz, 1966) illustrate slightly improved segmentation performance. Nevertheless, the computational cost is doubled in this process.

- Mixture of Gaussian model

Another technique used to improve the discriminative power of TGMRF features is mixture of Gaussian modelling. Peng et al. (2005) propose to model a mosaic texture image that composed of more than one texture class using a mixture of Gaussian (MOG) distribution. Peng et al. (2005) model unifies GMRF model and MOG model for texture segmentation. Model parameters are estimated using Expectation Maximization (EM) algorithm. Four mosaic images created using Brodatz texture (Brodatz, 1966) are segmented using the proposed features and have achieved acceptable segmentations. However, further experiments have not been carried out to achieve generalized performance.

- Features derived based on TGMRF descriptors

Instead of using model parameters directly as features, features derived from the estimated model parameters are suggested by Cesmeli and Wang (2001). The proposed features behave as variances in four neighbour directions. Second order GMRF model is employed and nine mosaic images are employed in segmentation. These features derived based on TGMRF descriptors have been able to perform relatively better compared to TGMRF descriptors.

- Multi-resolution GMRF descriptors

The drawback in some of the improved GMRF based features are the parameter estimation associates iterative optimization methods, for example, simulated annealing, which is a stochastic relaxation technique and is computationally very demanding. One way of reducing this computational burden is using multi-resolution techniques (Krishnamachari and Chellappa, 1997). Computational complexity is reduced because much of the work is done at coarse resolutions, where there are significantly fewer pixels to process. Also examining image in coarser resolution with a given smaller neighbourhood size also implies decisions are based on much larger neighbourhoods, without increasing computational complexity which would result by using larger neighbourhoods at original image resolution. It also leads the way to utilizing information at various scales and relax the problem of model selection (Comer and Delp, 1999). Krishnamachari and Chellappa (1997) propose multi-resolution GMRF with application to texture segmentation. In this method a given image at the fine resolution is modelled by a GMRF and coarser resolution images are obtained by sub sampling the fine resolution data and modelled by approximated GMRF models. There are a number of other Multi-resolution GMRF techniques reported in the literature (Comer and Delp, 1999).

The improved methods of GMRF features discussed here achieve relatively better performance compared to TGMRF features. However, the price paid for improving the performance is higher computational cost and complexity in the proposed algorithms. Nevertheless, small scale datasets are associated with the experiments and the generalized performance is not properly inspected. Despite the introduction of improved GMRF features, the original TGMRF features have been more commonly employed in recent GMRF literature. Therefore, current research examines TGMRF features in depth and focus on improving TGMRF descriptors using other means which are simple and effective. Large scale evaluations on descriptive power are conducted to bring up generalized conclusions and well-establish the improved descriptors.

3.5 Issues of TGMRF Descriptors

The traditional model parameter estimation process of GMRF can be considered as a global parameter estimation scheme. This is because the estimated parameter vector

tries to characterize the global texture appearance rather than the local structures of the texture. The parameter estimation process usually depends on a fairly large estimation window capturing the complete homogeneous texture. Relatively large sample sizes are achieved by sampling the estimation window. These samples carry the local structural information about the texture, however the global parameter estimation process results in more generalized estimates that tend to over-smooth the local information about a texture. The Gaussian notion and global parameter estimation in GMRFs leads to over smoothed models of texture, specially deterministic texture. The parameters of over-smoothed models have a lower discrimination power as texture features. On the other hand, the GMRF model selection further affects the quality of the GMRF parameter estimates which are employed as texture features. We will look at these aspects in detail in the following subsections.

3.5.1 Model Checking

GMRF is widely popular because it avoids the difficulties in parameter estimation and therefore makes the process analytically and computationally efficient (Besag and Kooperberg, 1995; Rue and Held, 2005; Manjunath and Chellappa, 1991). Using a Gaussian model is a valid assumption for texture modelling because most of the real textures obey the Gaussian notion (Manjunath and Chellappa, 1991). GMRF based features produce good results for homogeneous, fine, stochastic textures, but poorly perform when characterizing more structured and macro textures (Petrou and Sevilla, 2006; Bosnjak et al., 1998).

The main reasons for the poor performance is the parameter smoothing caused by under-fitted GMRF models. This causes loosing of much information contributing to the better description of the texture, specially the local texture characteristics. The Gaussian assumption, which leads to linear models, itself leads to an under-fitted model with the large sample size extracted from a larger estimation window. To elaborate this fact here we consider a model checking process.

The parameter estimation with LSE in the GMRF, which is defined by the Gaussian form and the linear dependency on neighbours, can be considered as a multiple linear regression problem. To assess how well the regression model describes the relationship between the neighbour pixels (explanatory variables) and a considered pixel (dependent variable) we use the analysis of residuals. The residuals represent the variations in the data that are not explained by the model. If the residuals are approximately normally distributed and have a constant variance it generally implies that the model is well fitted to the data.

Figure 3.3 shows the scatter plots of the standardized residuals against the standardized estimated value ($\hat{y}_s$), for four textures. The main observation that can be seen from

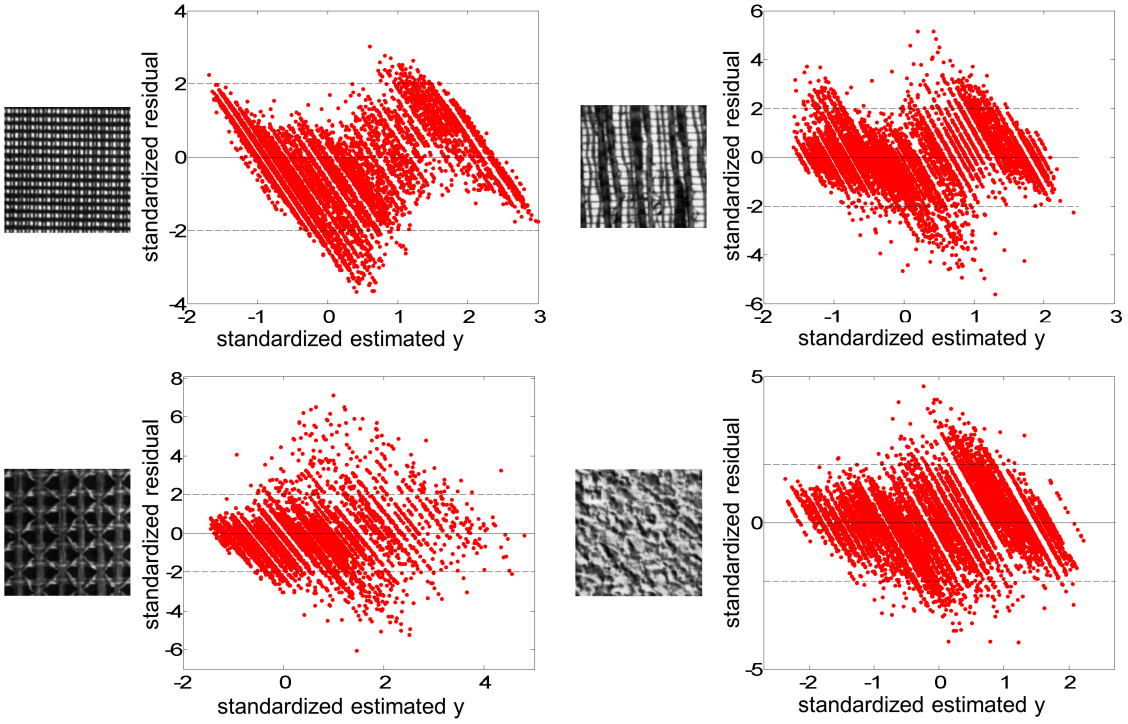


Figure 3.3: Scatter plots of standardized residuals of residual $(y_s - \hat{y}_s)$ for four textures. 95% confidence interval is given by dash lines. The extracted TGMRF feature vector $\mathbf{f}$ is used to calculate $\hat{y}_s$ at each pixel. $n = 3$ is used.

scatter plots of Figure 3.3 is that the residuals have not properly followed a normal distribution. To satisfy the normal condition, the residuals should be randomly spread in a band clustered around the horizontal line through zero. The spread is not totally random in the scatter plots appear in Figure 3.3 and have many outliers beyond the 95% confidence interval (Marked in dash lines). To have a constant variance, the residuals should exhibit no pattern and there should be no observed relationship in the graph. However, the residual plots in Figure 3.3 clearly do not have a constant variance (heteroscedasticity) as they have differnt variations with different $\hat{y}_s$. These facts indicate that the linear model fit to the data is not a proper fit and thus may not model the texture properly. Hence the Gaussian and the linear assumption may lead to estimates that characterize under-fitted models. This kind of parameter smoothing could cause the loss of important structural information about the texture.

3.5.2 Model Selection and Smoothing Effect

The previous subsection explained how the Gaussian form and the linear dependency in TGMRF texture modelling lead to over smooth models which negatively effect the texture feature extraction process. Further to this fact, model selection in the traditional GMRF also contributes to inaccurate parameter estimates which ultimately results in poor texture features.

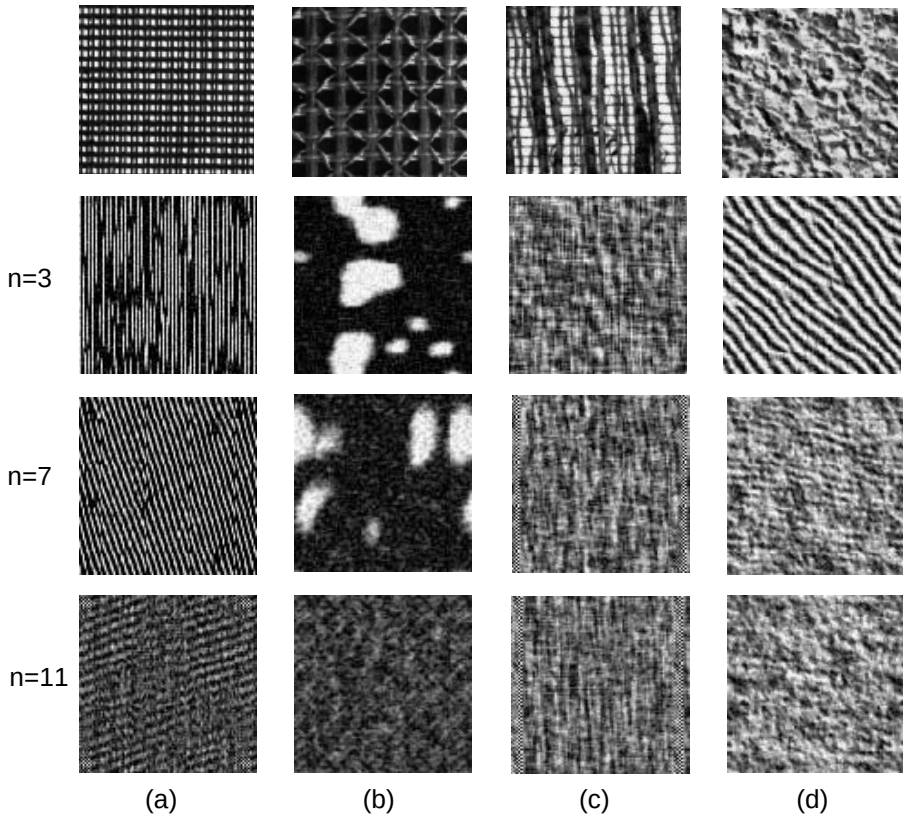


Figure 3.4: Synthesized images using TGMRF feature vector $\mathbf{f}$ achieved by employing different model orders.

Model selection in this context refers to the neighbourhood size (model order) selection in GMRFs. In TGMRF feature extraction, the analysis of spatial interactions is limited to a relatively small neighbourhood. i.e. the usage of small neighbourhood sizes or low model orders (Liu and Wang, 2003). If the model order is freely increased to follow the pattern size, the number of interaction parameters in the model increases quadratically. Such an increase in the model parameters leads to a computationally more expensive estimation process. During the parameter estimation however, model order should be approximately equal to the pattern size. Such a model order preserves the Markovianity (Stan et al., 2002). But the pattern size is usually unknown and the selected small model orders may not always be adequate to properly characterize the texture.

These difficulties have been reported before however, many studies in the literature have been persuaded to choose manually fixed small neighbourhood sizes (Manjunath and Chellappa, 1991; Chellappa and Chatterjee, 1985; Cesmeli and Wang, 2001; Mahmoodi and Gunn, 2011; Deng and Clausi, 2004). As a result, the adequacy of these features to characterize textures of various types and pattern scales is rarely checked (Liu and Wang, 2003).

In Figure 3.4 some examples of synthesized images using GMRF parameters are illustrated. A Gibbs sampler is used to synthesize the images from an initial Gaussian noise

image (Li, 2009)(see Appendix B). The synthesized image will be considered as the estimated image ($\hat{Y}$). The true original images (Y) are shown in the first row and the model parameters are estimated with different neighbourhood sizes. The estimated model parameters from different size neighbourhoods are then used to synthesize the new texture images.

By going down along each column of Figure 3.4 we observe that when the neighbourhood size, n increases the synthesized image gets more similar to the original. Actually there exists a correct value of n which provides better synthesised images (Petrou and Sevilla, 2006). However, the synthesized image even using a larger neighbourhood size ($n = 11$) does not look similar to the original (for example Figure 3.4b). Also deterministic and large textures are more likely to be poorly modelled than fine and stochastic texture (Figure 3.4a,b). More examples of synthesized images similar to the ones in Figure 3.4 generated by TGMRF parameters can be found in Petrou and Sevilla (2006). Thus, the improper model selection also have an impact on the ability of the parameter estimates to describe the texture.

3.6 Conclusions

GMRFs have been widely selected over MRFs due to its well defined model form with finite number of model parameters which can be efficiently estimated. Parameter estimation of GMRF conditional model is achieved using LSE and the model parameter estimates have been used as texture features. These features have performed successfully in small scale texture discrimination tasks and application specific tasks. However, the global parameter estimation process involving large estimation windows results in highly biased model parameter estimates which have poor texture discriminative power in texture segmentation and classification. The Gaussian notion and the linear model assumption in GMRFs lead to over smoothed models of texture. The residual analysis of GMRFs applied to texture suggests that most of the variations and non-linearities in the complex structure of texture are not properly characterized by GMRF model parameter estimates. Also there is no specific way introduced in the literature for selecting the model order of GMRFs. Usually lower model orders are assumed to lower the computational cost and to avoid curse of dimensionality. However, this further affects the quality of GMRF parameter estimates. Therefore, both factors, i.e. the Gaussian formulation with the linear dependency and the choice of lower model orders, cause smoothed estimates of model parameters which poorly model the texture. This affects negatively in TGMRF texture feature extraction because the estimates of model parameters are explicitly used as texture features. However, methods of overcoming these problems are rarely investigated. The solution introduce in this research, which will be discussed in

the next chapter, uses a flexible non-parametric GMRF based method to capture complicated data structures in texture rather than using the model dependent parametric GMRFs.

Chapter 4

Local Parameter Histograms (LPH) Descriptor

4.1 Introduction

The GMRF parameter smoothing discussed in the previous section is a major problem that leads to loss of important structural information of the texture. Especially when the texture is deterministic and structured (Petrou and Sevilla, 2006). In the texture feature extraction domain this is a significant drawback in constructing discriminative texture features. Comparative studies have shown that the GMRF texture features inherit a reduced discriminative ability in large scale empirical texture evaluations, including many deterministic textures (Ojala et al., 2001; Hadjidemetriou et al., 2003; Pietikäinen et al., 2000; Petrou and Sevilla, 2006).

Some of the methods in GMRF based literature which have been proposed to enhance the discriminative power of traditional GMRF features were discussed in the Chapter 3. However, these techniques are mostly application oriented and have not been used to perform large scale texture classification or segmentation to illustrate the generalized performances.

On the other hand, drifting away from the GMRF community for texture analysis, more solutions addressing the linear regression problem have been discussed in other fields of research. Specially in the economic and finance community, methods such as the autoregressive conditional heteroskedasticity (ARCH) model have been introduced which can deal with the non constant variance (Engle, 2001). Moreover, other alternative techniques such as non-linear regression can also be used.

In this research we aim to preserve the simplicity of GMRF parameter estimation and to maintain the resemblance with the GMRF feature extraction process as much as possible when finding a new solution to the problem. The main objective is to harness

more discriminative texture features based on GMRFs in the texture feature extraction process. To find a solution to the GMRF smoothing problem, we investigate the performance of locally estimated GMRF features based on local linear regression. i.e. by using a localized estimation process rather than a global parameter estimation process.

Local regression, also referred to as the kernel regression, is a non-parametric method that depends on data itself rather than relying on a specific pre-selected model (Hastie et al., 2013). This framework gives a rich mechanism for computing point-wise estimates with minimal assumptions about the global model. For the local regression the underlying model may remain totally unspecified. The local linear regression, a variant of local regression, fits many localized linear models to describe any signal (Hastie et al., 2013). Here, we use the local linear regression to simplify the estimation process over the local regression and to maintain the direct link to the GMRF parameter estimation.

The concept of localized parameter estimation has been used as an effective tool for image de-noising, interpolation and other image processing tasks (Takeda et al., 2007; Zhu et al., 2012; Gupta et al., 2008; Portilla et al., 2003). However, the local parameter estimation in the GMRF based texture feature formulation has not been addressed before.

Furthermore, recent studies on texture and object recognition have demonstrated that image representation based on distributions of local features are surprisingly effective (Zhang et al., 2007; Pietikäinen et al., 2000). Distributions of local features such as local binary patterns, spectral histograms and non parametric MRF methods have demonstrated impressive results in texture classification and segmentation (Ojala et al., 2002; Liu and Wang, 2003; Zhang et al., 2007; Ojala et al., 2001; Pietikäinen et al., 2000).

Following this in the present study, we investigate the performance of distributions of local parameter estimates. The distributions of local parameter estimates capture the local dependencies of the texture through the localized parameter estimation stage and the global appearance through the histogram construction stage. Also locally estimated models are partial representatives of the texel or texture patterns. Hence, they are not only enriched with spatial interaction information, they also carry structural information about the texture primitives. Furthermore, histogram construction captures the statistical properties about the distribution of primitive characteristics defined by local models.

Small neighbourhood sizes are generally more favourable for constructing local feature distributions. This is because it enables fitting localized models easily into a localized area. Therefore, local parameter estimation is suitable when small neighbourhood sizes are involved and hence relax the requirement for proper model selection.

4.2 Local Linear Regression

The benefits of a GMRF model are its simplicity and ease of parameter estimation. GMRF represents the conditional expectation via a linear model which is the weighted sum of neighbour values (see equation (3.4)). The parameter estimation procedure therefore becomes a linear least square regression problem where analytical solutions simply exists.

However, in the case where a linear regression function does not capture the underlying true model of data, estimates are biased and are over-smoothed leading to poor results (Gupta et al., 2008). This happens in texture modelling with GMRFs, specially where texture inside the estimation window is more structured and deterministic (Petrou and Sevilla, 2006). This is a major drawback in GMRF linear regression parameter estimation.

As a solution, a localized version of the model fitting can be employed by using local linear regression technique, a non parametric method with more flexibility in estimating regression functions (Hastie et al., 2013). Local linear regression exploits the fact that any sufficiently smooth arbitrary function can be well approximated by locally straight lines or hyper-planes (in higher dimensions) fitted over a small enough local subset of the domain (Gupta et al., 2008). i.e. relaxation from a globally linear model to one that is locally linear. The local linear regression closely models the underlying function by fitting a different but simple model separately at each pixel. This method belongs to a category of regression techniques known as kernel smoothing techniques (Hastie et al., 2013). Here, the kernels are mostly used as a device for localization.

In GMRF feature extraction, the difference between linear regression and the local linear regression (sometimes referred to as small model estimation here) is the selection of the estimation window size. This is explained by an example in Figure 4.1 for a model with one predictor. The continuous blue line shows the linear regression fit and dashed red line illustrate the local linear regression fit. It can be clearly seen from Figure 4.1 that the coefficients of ordinary linear regression represent an over-smoothed model. In a texture feature extraction task this is a great disadvantage because it misses out some important structural information about the texture. Unlike in a synthesis problem, discriminative features are required rather than reliable modelling of the texture in classification and segmentation problems. Therefore, in this study we explore the performance difference that can be achieved by non-parametric localized linear regression models.

4.2.1 GMRF Small Model Estimation

We name the process of localized parameter estimation using local linear regression as *small model estimation* because the individual models are constructed based on a smaller

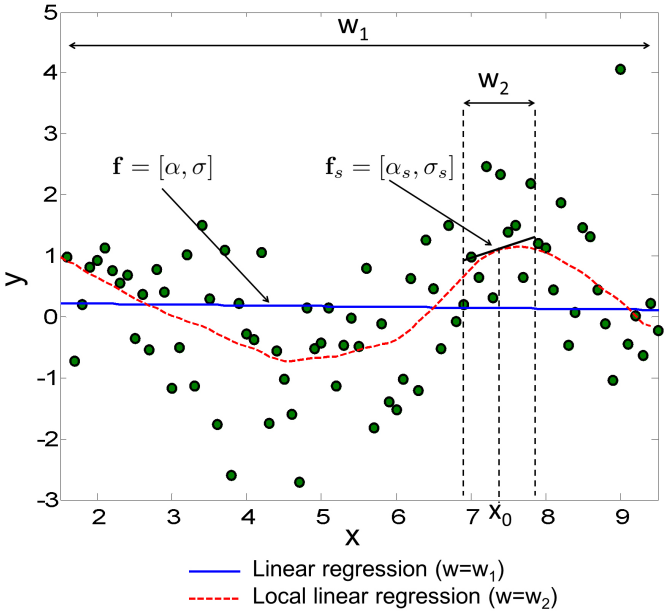


Figure 4.1: Linear regression and local linear regression model fitting. w - estimation window size. In local linear regression estimated value of y at x_0 is found by fitting a local linear model at x_0 using the samples inside w_2.

localized area. Here, rather than fitting a linear model to the entire set of observations in Ω, local linear regression fits a simple model to only a small subset of observations in a region Ω_s local to each pixel minimizing the local error,

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{s}} = \arg\min_{\boldsymbol{\alpha_s}} \sum_{t \in \Omega_s} \epsilon_t^2 = \arg\min_{\boldsymbol{\alpha_s}} \sum_{t \in \Omega_s} \left(y_t - \boldsymbol{\alpha_s}^T \bar{\boldsymbol{y}}_t \right)^2 \tag{4.1}$$

leading to,

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{s}} = \left[\sum_{t \in \Omega_s} \bar{\boldsymbol{y}}_t \bar{\boldsymbol{y}}_t^T \right]^{-1} \left[\sum_{t \in \Omega_s} \bar{\boldsymbol{y}}_t y_t \right] \tag{4.2}$$

and the variance parameter of the model,

$$\hat{\sigma}_s^2 = \frac{1}{|\Omega_s|} \sum_{t \in \Omega_s} \left(y_t - \hat{\boldsymbol{\alpha}}_{\boldsymbol{s}}^T \bar{\boldsymbol{y}}_t \right)^2 \tag{4.3}$$

The estimation window is now Ω_s instead of Ω. When Ω_s approaches Ω, TGMRF parameters can be achieved and parameter dependence on the linear assumption increases. The adaptive TGMRF parameter estimation can be easily adjusted to local linear regression parameter estimation by reducing the estimation window size, w while addressing computational inconsistencies that can arise by smaller estimation windows. Note that this is equivalent to using a uniform kernel or in other words a weighting function where samples inside Ω_s are weighted by a factor one and samples outside Ω_s are weighted by a factor zero (Hastie et al., 2013).

From the viewpoint of a homogeneous texture, TGMRF feature extraction assumes that the linear regression coefficients are constant over a homogeneous texture. However, parameters estimated at each pixel by local linear regression have a higher spatial variation and suggest that there are multiple spatially adjacent distinct linear models contributing to form a texture.

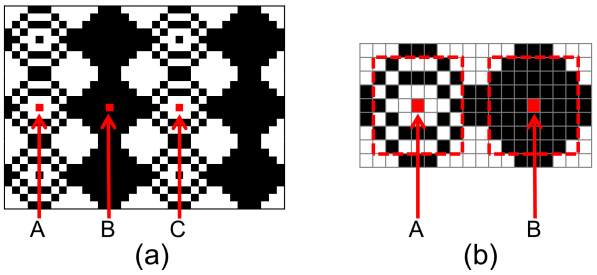


Figure 4.2: (a) Synthetic deterministic texture (b) texel of the texture in (a) and two n=7 neighbourhoods posing different interactions on their centre pixels.

Furthermore, the spatial variations in parameter estimates are much higher when a selected model order, n is inadequate to describe the texture. For example, in Figure 4.2 the chosen neighbourhood size, n is smaller than the actual texel size. The pixel A and pixel B have different interactions with their corresponding $n = 7$ neighbourhoods (Figure 4.2b). For that reason interaction parameters should be able to have some degree of spatial variations when low order GMRFs are involved. In such a situation the localized modelling can better describe the underlying process which minimizes the smoothing incurred by parameter estimation with larger sample sizes. Localized parameter estimation overcomes the strong dependence on choice of the model order, n. It enables low order GMRF models more suitable for texture description preserving spatial variations, spatially when relatively large texture patterns are defined by lower model orders.

The model parameter vector $\mathbf{f}_s = [\hat{\boldsymbol{\alpha}}_{\boldsymbol{s}}{}^T, \hat{\sigma}_s]^T$ at a pixel characterizes the locally fitted GMRF model simply referred to as the small model. These local parameters can better describe any complex texture compared to the TGMRF global features.

4.3 Concerns in Small Model Estimation

4.3.1 Estimation Window Size

The size of the region Ω_s plays a significant role in the local estimation results. The region Ω_s corresponds to the estimation window of size $w \times w$ pixels or a uniform kernel with weights equal to one. In our study we suggest smaller sizes of the estimation window w to achieve the localized estimation process. However, smaller estimation window size provides fewer samples for the estimation process. Therefore, the issues arising from small sample sizes in the estimation process should be carefully considered.

First and foremost we need to have a sample size at least greater than the number of interaction parameters (explanatory variables). When the number of samples is less than the number of model parameters, an infinite number of solutions can exist. This is because when the number of samples are equal to the number of independent variables the observation space can be fully explained and any additional variables will cause variable redundancy. Usually the generalized inverse could be useful in such a situation which results in a unique solution which is the least squared solution as well as the minimum norm solution (Sullivan and Liu, 1984; Penrose, 1956). However, the generalized inverse is relatively expensive.

Therefore, when constructing the texture features, despite the fact that we need to localize the estimation process, we make sure the estimation window size is large enough to provide the required number of samples to obtain unique solutions in the parameter estimation. A heuristic rule is used to select the estimation window size which makes sure the number of samples are always more than the number of model parameters. In GMRF model, the number of interaction parameters depends on the size of the neighbourhood and is equal to $(n^2-1)/2$. Since the samples are extracted from a region inside the $w \times w$ estimation window, the number of overlapping samples that can be sampled from the estimation window is equal to $(w-n+1)^2$. Thus, we deduce a value for w as follows.

The number of samples must be greater than or equal to the number of interaction parameters. Therefore,

$$(w-n+1)^2 \geq (n^2-1)/2 \tag{4.4}$$

Next we select the case which is well above the equality criteria, that is the number of samples is equal to n^2.

$$n^2 \geq (n^2-1)/2 \tag{4.5}$$

This gives roughly twice the number of samples than the number of interaction parameters. If this is the case, the number of samples is equal to n^2, which is,

$$(w-n+1)^2 = n^2 \tag{4.6}$$

leading to the value of the estimation window w,

$$w = 2n-1 \tag{4.7}$$

which is small enough to fit localized models while providing unique solutions to the least squares problem. The value for size w selection is illustrated in Fig 4.3. Therefore, in this study for local linear regression the estimation window size is selected as $w = 2n-1$.

The size of w is a function of the neighbourhood size n and thus we only need to specify n in the parameter estimation process.

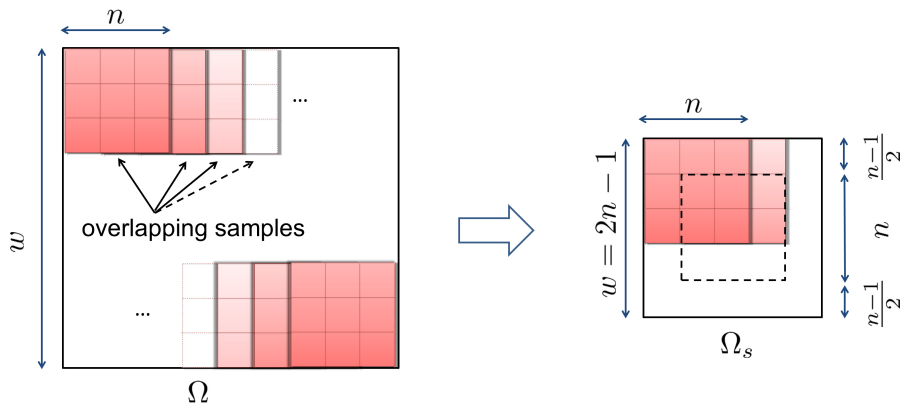


Figure 4.3: Estimation window size selection. The estimation window size become $w = 2n - 1$ when number of samples extracted from it is equal to n^2.

4.3.2 Handling Ill Posed Inverse Problem

When reducing the size of the estimation window, the number of samples for the estimation process becomes limited. This situation further causes two major related problems. One is that the estimation process can become inconsistent or an ill-posed problem because the matrix inversion involved with the estimation of interaction parameters (see equation (4.2)) can become non invertible. This non-invertiblity issue can occur when several of the explanatory variables (here the neighbour values) have dependency on each other. The other major problem is that too few training samples can result in regressions with incorrectly fitted models when outliers exist giving unacceptable large values for the estimates.

The ill-conditioned problem has been extensively studied by statisticians in the domain of approximation theory (Björkström, 2001). Several techniques have been suggested to overcome this problem and it has been shown that in many practical contexts, Tikhonov regularization, also know as the ridge regression perform well in finding approximate solutions to ill-conditioned problem (Björkström, 2001; Hastie et al., 2013). Tikhonov regularization minimizes both the squared error term plus the Euclidean norm of the interaction parameter vector $\boldsymbol{\alpha_s}$.

Therefore, the Tikhonov regularization is applied as below to regularize the local error minimization process and estimate model parameters at pixel s.

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{s}} = \arg\min_{\boldsymbol{\alpha_s}} \sum_{t \in \Omega_s} \left(y_t - \boldsymbol{\alpha_s}^T \bar{\boldsymbol{y}}_t \right)^2 + c^2 \boldsymbol{\alpha_s}^T \boldsymbol{\alpha_s} \tag{4.8}$$

where c is a constant and is called the regularization parameter which controls the trade-off between minimizing the error and penalizing the magnitude of the parameters. The regularized local linear regression estimation results in,

$$\hat{\boldsymbol{\alpha}}_{\boldsymbol{s}} = \left[\sum_{t\in\Omega_s} \bar{\boldsymbol{y}}_t \bar{\boldsymbol{y}}_t^T + c^2 \boldsymbol{I}\right]^{-1} \left[\sum_{t\in\Omega_s} \bar{\boldsymbol{y}}_t y_t\right] \tag{4.9}$$

where $\boldsymbol{I}$ is the identity matrix. From equation (4.9) it can clearly observed that in Tikhonov regularization a value c^2 is added to the elements on the diagonal of the matrix to be inverted making it a diagonal dominant matrix. The diagonally dominant matrices are non singular and therefore the non-invertiblity issue is solved in the case that the matrix inversion in equation (4.2) is ill-conditioned (Björkström, 2001). Also it can be seen from equation (4.8) that because the euclidean norm of $\boldsymbol{\alpha}_{\boldsymbol{s}}$ is also minimized it reduces the values of the estimates of interaction parameters towards zero appropriately (given c is chosen correctly). This helps to overcome the over-fitting problem when outliers exist in the selected sample and avoid unacceptable large estimate values. The value for c is selected experimentally.

By ensuring the number of samples is greater than the number of model parameters (w selection) and by applying Tikhonov regularization, we can achieve a well posed unique localized parameter estimation process using local linear regression.

After estimating the interaction parameter vector in this way, next the variance parameter is obtained according to equation (4.3). The entire least square estimation process with the w selection and the regularization is referred by the name *small model estimation* as introduced before. In a localized parameter estimation environment the small model estimation is suitable for estimating local parameters.

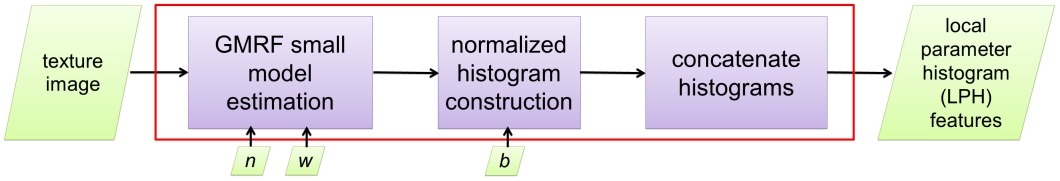


Figure 4.4: Local parameter histogram construction and associated variables.

4.4 Local Parameter Distributions

Distributions of locally extracted features are a popular choice in many texture feature extraction techniques (Ojala et al., 2002; Liu and Wang, 2003; Zhang et al., 2007). Small model estimation provides local texture features describing local interactions and local structures of a texture. Therefore, the distributions of local parameters achieved by small model estimation is employed as a novel improved texture descriptor. These features are named as local parameter histograms (LPH) descriptors.

The small model estimation is performed over each pixel of the given texture image by the sliding window estimation technique. The estimation of small models at each pixel results in a parameter vector at each pixel describing the local dependencies. If one of the parameters from the parameter vector is considered in the spatial domain, a parameter image, F_j can be defined as $F_j = \{f_s(j)|s \in \Omega\}$ where $f_s(j) \in \mathbf{f_s}$. Therefore, for all the parameters in $\mathbf{f_s}$ we get separate parameter images leading to a parameter image stack $F = \{F_j|j = 1, \ldots, (n^2+1)/2\}$ (see Figure 4.5). Note that $(n^2+1)/2$ is the number of model parameters in $\mathbf{f_s}$ including the variance parameter.

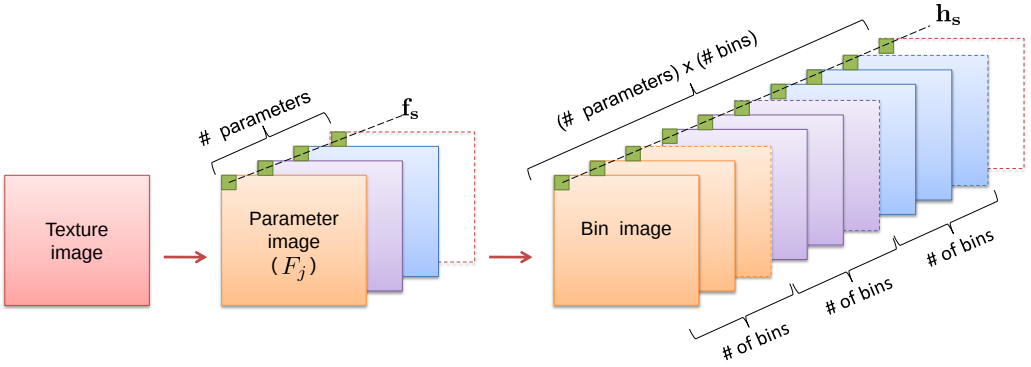


Figure 4.5: Parameter images and bin images result from local parameter estimation and local histogram construction respectively. For the local parameter estimation, a sliding window of size w and for the histogram calculation a sliding window of size b are employed.

The local parameter estimates on a parameter image, F_j contain more spatial variations than adaptive TGMRF parameters. A few sample parameter images are shown in Figure 4.6 for a mosaic image. In these parameter images we observe spatial repetitiveness of local parameters according to the repetitive pattern of a given homogeneous texture (Figure 4.6).

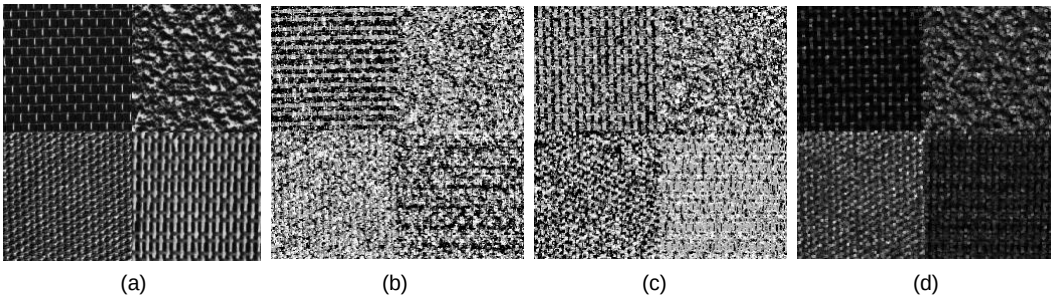


Figure 4.6: Parameter images. (a) A mosaic texture image, (b)-(d) parameter image corresponding to horizontal interaction parameter, vertical interaction parameter and variance parameter, respectively.

The texture segmentation problem involves mosaic texture images comprising many textures. For texture segmentation purposes, the texture features at each pixel are required. Therefore, on each parameter image, the normalized parameter histograms,

$h_s(j)$ are constructed by sliding another window of size $b \times b$ pixels to formulate the local histogram at each pixel. The concatenated LPH descriptor at pixel s is then given by,

$$\mathbf{h}_s = [h_s(1)^T, \ldots, h_s(j)^T, \ldots, h_s((n^2+1)/2)^T]^T \tag{4.10}$$

The local parameter histogram construction is graphically illustrated in Figure 4.4 and the resulting parameter and bin images are explained in Figure 4.5. In segmentation, the choice of histogram calculation window size ideally should be roughly equal to or be greater than the texel size of the largest pattern in the texture. This will lead to a perfect homogeneous texture descriptor. Typically, size b is selected independent of the size w. Furthermore, b should be large enough to provide a sufficient number of samples (equals to b^2), given the number of bins, in order to construct a proper histogram. Frequently used rule of thumb for constructing histograms is that the number of samples should be roughly equal to or be greater than the square of number of bins ($b^2 \geq bins^2$). Therefore, once the number of bins is known, the value of b should be roughly the number of bins or greater than that to achieve a good performance. However, the value of b cannot be increased unnecessarily in texture segmentation because of the increase in the boundary localization error. In this study the effect of b on segmentation error is experimentally analysed in section 6.3.3 and a range of suitable b values are suggested. Therefore, the value of b used in this work is roughly equal to the number of bins or higher in texture segmentation experiments.

On the other hand, texture classification involves images comprising one specific texture. For texture image classification, features at each pixel is not required. Thus, the histogram of the whole parameter image F_j is directly used as the LPH descriptor. i.e. b = imageSize is used for constructing the distributions giving one concatenated histogram $\mathbf{h}$ per image.

The LPH construction is illustrated in Figure 4.4. The number of bins is manually fixed. For example, if $n = 3$ is selected, there are five different model parameters governing the GMRF model, i.e. $(n^2-1)/2 = 4$, four interaction parameters and the variance parameter. If 10 bins for the histogram is employed, then there will be a 50 dimensional feature vector.

In conclusion, for the construction of LPH descriptors, a predefined neighbourhood size n for GMRF representation and a histogram calculation window size b for constructing local histograms should be predefined. For a texture classification task histogram calculation window size b is same as the image size.

4.4.1 Effect of Local Parameters on LPH descriptors

When the estimation window size w is smaller, estimated local parameters have more spatial variations visible on the parameter images. This gives a higher information content to the histograms of local parameter estimates. However, to meet the consistency criteria we have bound the minimum size of w to $w = 2n - 1$ which depends on the model order, n, as explained in section 4.3.

In Figure 4.7, the entropy of the local parameter histogram with changing estimation window size is shown. Here, the entropy of the histogram depicts its ability to capture spatial variations in the parameter estimates associated with the texture. Higher entropy values suggest the presence of more spatial variations and vice versa. When w increases, entropy gradually drops implying that spatial variations are progressively smoothed out by the estimation process. Therefore, smaller estimation windows are also preferable to construct local feature distributions.

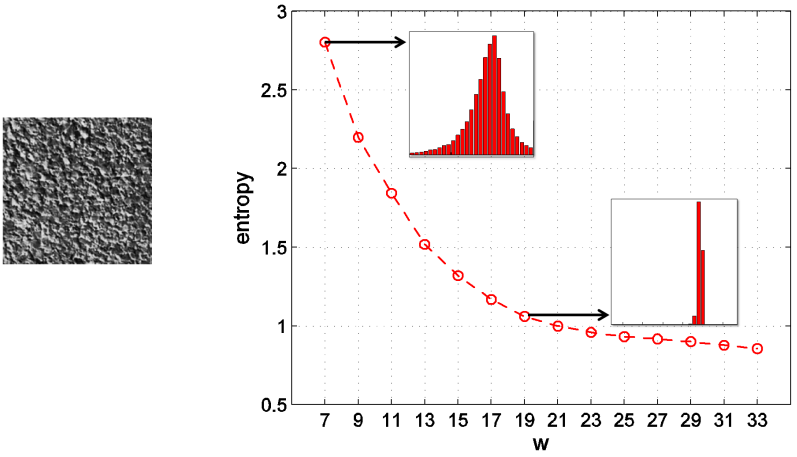


Figure 4.7: Entropy of the histogram of horizontal interaction parameter of the given texture image in top right corner with changing estimation window size w ($n = 5$ is used).

4.5 Comparison to TGMRF Features

LPH features consider the distributions of GMRF local parameters achieved using a comparatively smaller estimation window to that of adaptive TGMRF features. However, these modifications can lead to a significant performance improvement. This is because the smoothing effect is reduced by the local model fitting and important structural information are preserved within the parameter estimate space.

Figure 4.8 illustrates LPH descriptors of horizontal and vertical interaction parameters and the variance parameter of GMRF model belonging to two different homogeneous

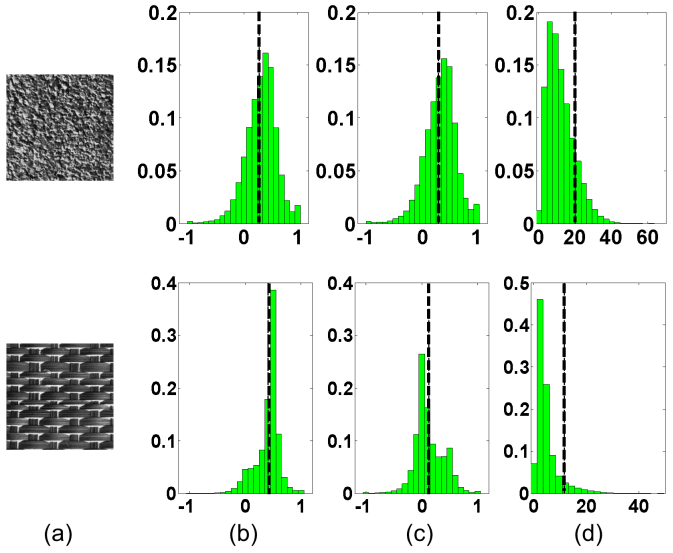


Figure 4.8: LPH descriptors of vertical, horizontal and variance parameters. (a) original images, (b) horizontal interaction parameter histogram, (c) vertical interaction parameter histogram, (d) variance parameter histogram, Dash lines represent parameter values estimated by adaptive GMRF method ($n = 5$ is used).

textures. The estimated parameter values from the TGMRF method are shown by the dashed lines in Figure 4.8. The TGMRF features have small differences between the two textures even though the displayed textures have large visual dissimilarities. In contrast, LPH descriptors demonstrate a significant difference in their distributions.

Figure 4.9 presents the pixel-wise adaptive features extracted from both methods for two sample mosaic images. The dimensions of the original feature vectors are reduced to two dimensions using principal component analysis (PCA) for illustration purposes. Each data point in the Figure 4.9 represents the reduced-dimension feature vector at a pixel.

In the estimation process of constructing TGMRF descriptors, the texture contained in the estimation window should be homogeneous and therefore, large estimation windows are employed. However near texture boundaries, non-homogeneity can occur leading to inaccurate features which cause boundary localization problems, especially when large estimation window sizes are involved (Xia et al., 2006b; Dharmagunawardhana et al., 2012). This means that in TGMRF feature extraction process the boundary localization problem is more prominent. Therefore, TGMRF descriptors have a larger overlap and a smaller separability (inter-class distance) between the two texture classes compared to LPH descriptors (Figure 4.9b,c). Also LPH descriptors have more local information about the texture than the TGMRF features leading to a tighter dispersion in class features, conveying lower intra-class variation (Figure 4.9c). From Figure 4.9 it is observed that LPH descriptors are better descriptive features compared to TGMRF features.

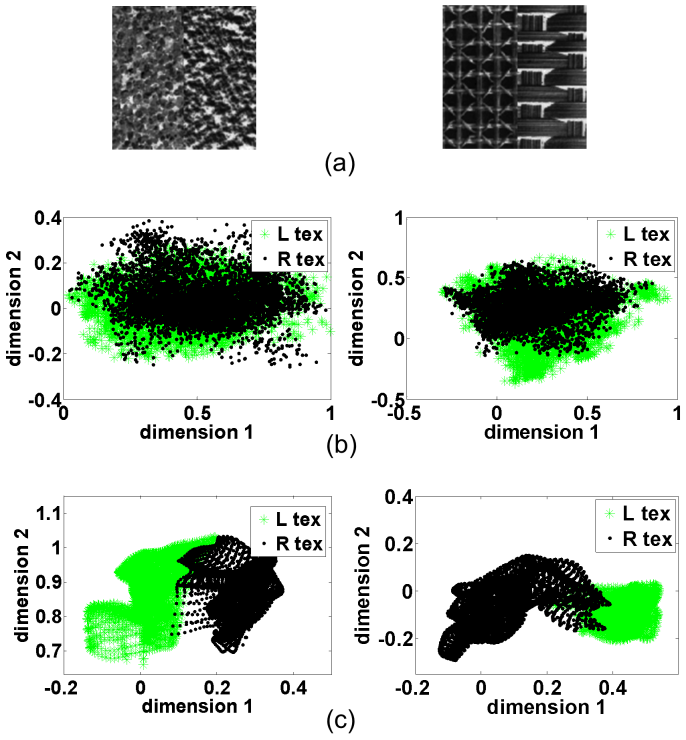


Figure 4.9: Feature vectors in 2D for two-texture mosaic images. (a) original images (b) texture features from adaptive TGMRF method ($n = 5, w = 21$), (c) texture features from adaptive LPH method ($n = 5, b = 21$). PCA is used to reduce dimensionality. L-tex = pixels belongs to left side texture and vice versa.

4.6 Conclusions

LPH descriptor is introduced in this chapter as an improved GMRF based texture feature. The construction of LPH descriptors has two main stages namely, a local parameter estimation process called small model estimation and a histogram construction stage. The small model estimation process estimates local model parameters using local linear regression which is a non-parametric method. Local linear regression is flexible in modelling complex data structures and has a lower bias to the model assumptions. The localized estimation process is achieved by selecting smaller estimation window sizes. An estimation window size selection criterion is proposed in this chapter based on the number of samples which can be obtained from the estimation window and the number of parameters in the model. The local estimation process encounters inconsistencies due to the small sample size used in the estimation process. These difficulties are overcome by using Tikhonov regularization. After the local parameter estimation, the histogram construction stage is carried out to formulate a homogeneous texture descriptor. The small model estimation captures the structural properties of the texture such as local interactions and local roughness of the texture. On the other hand, histogram construction captures statistical properties of the texture such as distributions of local features.

Compared to the TGMRF features, LPH features can produce a more discriminative reliable texture description.

Chapter 5

Rotation Invariant Descriptors

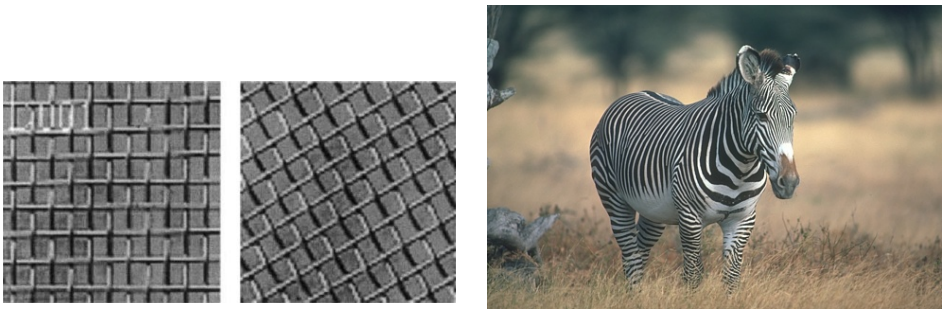
The LPH descriptors can be considered as scale invariant features up to some degree due to the histogram construction. This is due to the scale invariance property of histogram features. However, they are not rotation invariant. This chapter considers techniques to achieving rotation invariant texture features based on LPH descriptors.

5.1 Background

The majority of existing texture analysis methods make the explicit or implicit assumption that the texture images are acquired from the same viewpoint, i.e. the same scale and orientation. Although the texture features are not view point invariant, these non-invariant texture features have been widely applied in texture analysis covering many fields including medical image processing and defect detection processes (Zhang and Tan, 2002).

However, in many other practical applications such as object recognition and image retrieval where more than one image is involved, it is very difficult or impossible to ensure that the associated texture components have the same translations, rotations or scaling across the image instances. Also from the perspective of natural image segmentation a specific texture component in a given image may have different physical surface rotations (see Figure 5.1). Based on the cognitive theory and our own perceptive experience, given a texture image, no matter how it is changed under translation, rotation and scaling, it is always perceived as the same texture image by a human observer (Zhang and Tan, 2002; Matthews et al., 2013). Therefore, invariant texture analysis is desirable for both the practical and experimental applications.

Over the last few decades an increased amount of attention has been given to invariant texture analysis, and several methods for achieving the rotation invariance have



(a) (b)

Figure 5.1: Encounters of rotation variations in textures. (a) rotated versions of texture images occuring in texture classification, (b) rotation variation encountered in texture segmentation (texture of zebra).

been proposed. In this subsection, some of the general techniques of achieving rotation invariance are reviewed under two categories.

5.1.1 Omnidirectional Texture Features

Omnidirectional texture features are constructed independent of the direction. The simplest approach of achieving rotation invariance in this way is by using invariant or isotropic statistics such as mean, variance and intensity histogram. However, their performance is poor because of the limited amount of information captured by them. Haralick et al. (1973) propose computing omnidirectional features from directional measures by averaging the GLCM based features over four directions. Mayorga and Ludeman (1994) employ isotropic texture edge statistics based on circular levels or neighbourhoods. In their study the circularly averaged differences in gray-level and the correlation along the circular levels are used as the texture features. The features extracted from filter responses achieved via isotropic filter kernels are also belong to rotational invariant omnidirectional features. They have been widely used in texture analysis providing higher texture classification rates (Porter and Canagarajah, 1997; Zhang et al., 2002).

Furthermore, model based approaches such as Circular Simultaneous Auto Regression (CSAR) simply average all the pixels on the unit circle neighbourhood into a single value associated with a single parameter, producing a model containing no directional information (Kashyap and Khotanzad, 1986). Mao and Jain (1992) extend this method to circular neighbourhoods with larger radii called, Multiresolution Rotation Invariant Simultaneous Auto Regression (MR-RISAR) model. An isotropic model parameter estimated from each circular level which represents the averaged neighbour interactions with the central pixel is used as the texture feature.

The problem with these approaches is that the directionality, an important characteristic of the texture, is lost when an isotropic feature is formulated. Thus, these features are more favourable with isotropic textures.

5.1.2 Rotation Invariant Directional Texture Features

This category of features includes texture features that measure directional characteristics of texture, yet are rotation invariant. A popular method of directional texture feature extraction is using a bank of oriented filters (Zhang and Tan, 2002). After obtaining the responses of directional filters, techniques such as maximum response (Ahonen and Pietikäinen, 2009; Varma and Zisserman, 2005), Fourier transform (Greenspan et al., 1994), polar transform (Haley and Manjunath, 1995), etc. have been used to achieve rotational invariant features.

Polar plots and polarogram is another approach based on polar transformation (use the polar coordinate system). Here the polarogram is a function of orientation. Therefore, a rotation in the original texture image results in a translation in polarogram. The features computed from the polarogram are invariant to rotation (Davis, 1981). Discrete Fourier transformation (DFT) is another popular choice of achieving rotation invariance. Because the magnitude of the DFT is invariant to translation, by performing DFT on a feature vector containing features from different orientations results in a rotation invariant feature. For example, Deng and Clausi (2004) use the DFT of the estimated interaction parameter vector that represents the influence of neighbours in different orientations which is extracted from An-isotiopic Circular GMRF (ACGMRF) model.

Furthermore, feature distributions of locally invariant features such as linear symmetric auto correlation measures, related covariance measures, rotation invariant local binary patterns and gray level difference have been successfully employed as rotation invariant features (Pietikäinen et al., 2000). The local features are made invariant based on neighbourhood operations such as circular shifting.

Unlike, omnidirectional features, these features preserve directional information. Nevertheless, they are generally computationally expensive than the omnidirectional features.

5.2 Achieving Rotation Invariance with LPH Descriptors

Two techniques belonging to previously discussed two categories of achieving rotation invariance are employed here. One way is by using circular shifted neighbours to achieve the rotation invariance and the other method is by using the Isotropic GMRF (IGMRF) model. Circular neighbourhoods are considered to construct rotation invariant features. A neighbourhood is defined by equally spaced neighbour pixel values located on a circle with a radius r. The neighbour values are achieved using bilinear interpolation similar to Ojala et al. (2002). The number of neighbours in a resolution is referred by p.

5.2.1 Circular Shifted Neighbour Method

The rotation invariant features ahieved by circular shifting of neighbour values is named as Rotation Invariant Local Parameter Histogram (RI-LPH) features.

A circular shifting process of neighbour values is performed on each neighbour vector extracted from the circular neighbourhoods. The neighbour pixel values in the neighbour vector $\tilde{\boldsymbol{y}}_s = col[y_{s+r}|r \in N]$ is circularly shifted based on the neighbour *difference* vector $\boldsymbol{d}_s = col[|y_{s+r} - y_s| \ \ |r \in N]$. The *difference* value is the absolute difference between a neighbour value and the considered centre pixel. The number of circular shifts to perform is calculated from the $\boldsymbol{d}_s$ by counting the shifts until the first element of the $\boldsymbol{d}_s$ vector becomes the maximum *difference* value. Once the number of circular shifts are figured out using $\boldsymbol{d}_s$ the neighbour vector of the considered pixel is circularly shifted by that amount. This process causes rotating the entire circular neighbourhood according to the direction of maximum *difference* value.This leads to a rotation invariant neighbour set $\boldsymbol{z}_s$ at location s. The algorithm for circular shifting neighbours is shown in Algorithm 5.2.1.

Algorithm 5.2.1: CIRCULAR SHIFTING NEIGHBOUR VALUES()

shftSize=1;
shftCount=0;
while $\boldsymbol{d}_s(1,1) \neq \max(\boldsymbol{d}_s)$
$\boldsymbol{d}_s = circShift(\boldsymbol{d}_s, \text{shftSize})$;
shftCount=shftCount+1;
end
$\boldsymbol{z}_s = circShift(\tilde{\boldsymbol{y}}_s, \text{shftCount})$

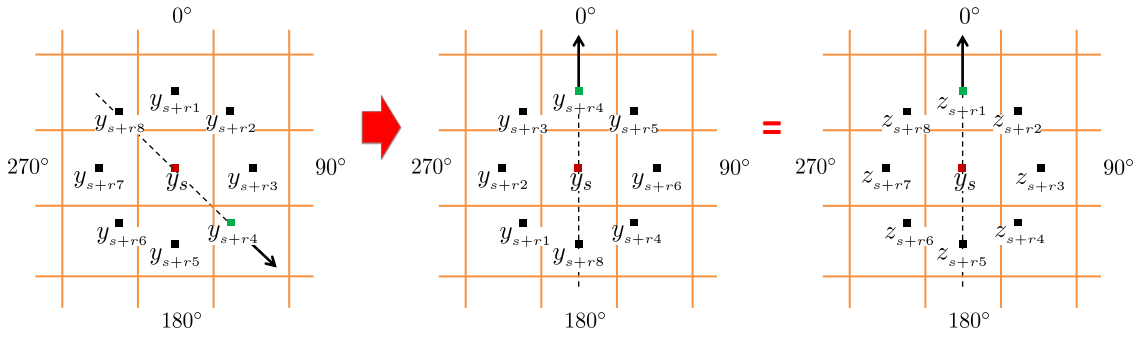


Figure 5.2: Rotation invariance by circular shifting.

The circular shifting process is graphically illustrated in Figure 5.2. Subsequently, usual localized estimation process is carried out as proposed in Chapter 4. It is rewritten here

Table 5.1: Different attributes associated in construction of LPH and RI-LPH descriptors. n- neighbourhood size, w-estimation window size, b-histogram calculation window size, r-radius of circular neighbourhood, p-number of neighbours in the circular neighbourhood.

attribute	LPH	RI-LPH
neighbourhood	square	circular
neighbourhood size	n	(r, p)
estimation window size	$w = 2n - 1$	$w = 4r + 1$
histogram calculation window size	b	b
rotation invariance	no	yes

with the notation of rotation invariant neighbour set.

$$\boldsymbol{\alpha_s} = \left[\sum_{s\in\Omega_s} \bar{\boldsymbol{z}}_s \bar{\boldsymbol{z}}_s^T + c^2 \boldsymbol{I}\right]^{-1} \left[\sum_{s\in\Omega_s} \bar{\boldsymbol{z}}_s y_s\right] \tag{5.1}$$

where $\bar{\boldsymbol{z}}_s = col[z_{s+r} + z_{s-r} | r \in \tilde{N}]$ according to the symmetric assumption and $\boldsymbol{I}$ is the identity matrix. The variance parameter of the model,

$$\sigma_s^2 = \frac{1}{|\Omega_s|} \sum_{s\in\Omega_s} \left(y_s - \boldsymbol{\alpha_s}^T \bar{\boldsymbol{z}}_s\right)^2 \tag{5.2}$$

The estimation window size w is selected similar to the LPH descriptors discussed in the Chapter 4. When r is given, the value of n can be written as $n = 2r + 1$ and therefore, $w = 2n - 1$ becomes $w = 4r + 1$ in terms of r. The differences between associated variables of constructing square neighbourhood based LPH descriptors and the circular neighbourhood based RI-LPH descriptors are shown in Table 5.1.

After performing localized parameter estimation, the histogram of each parameter image is constructed and concatenated to form the final feature vector. The process is graphically illustrated in Figure 5.3.

5.2.2 Isotropic GMRF Method

The second method we analyse here to achieve rotation invariance is by using IGMRF, called Isotropic Local Parameter Histogram (I-LPH) descriptors. IGMRF is a special case of GMRF also known as circular symmetric GMRF and further simplify the GMRF model (Kashyap and Khotanzad, 1986). IGMRF models the non directional isotropic textures in a simplified rotational invariant framework with only two model parameters. The parameter estimation is simple and fast compared to other MRF models because solutions for parameters can be found analytically and without requiring any matrix

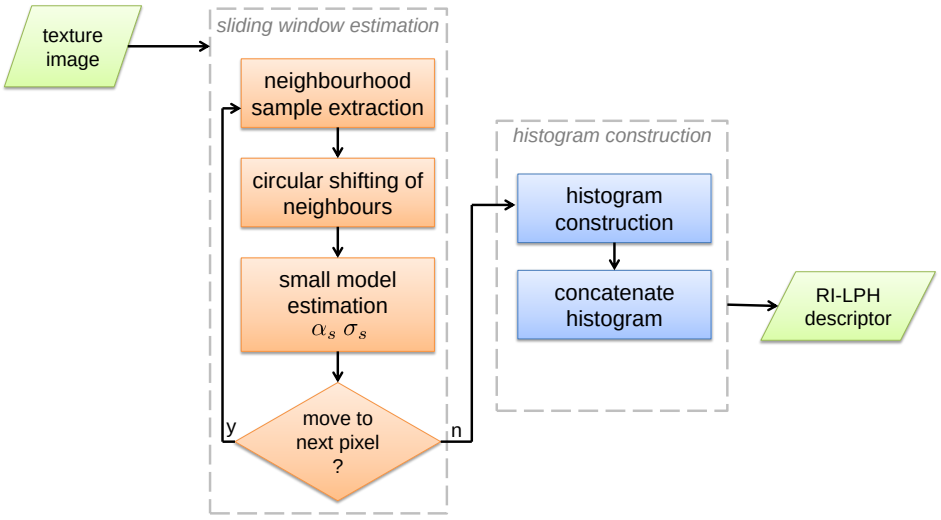


Figure 5.3: Construction of RI-LPH descriptors.

inversion. IGMRF model is given by,

$$p(y_s|y_{s+r}, r \in N; \alpha, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y_s - \alpha \sum_{r\in N} y_{s+r}\right)^2\right\} \tag{5.3}$$

where α and σ are the model parameters to be estimated. y_{s+r} are the neighbours in the circular neighbourhood of y_s. The circular neighbourhoods are employed here similar to RI-LPH descriptors defined by (r, p). At each pixel, the localized estimation procedure is carried out similar to RI-LPH descriptors with the estimation window size $w = 4r+1$.

LSE is used here for estimating α parameter of the model. As discussed before the main assumption behind the LSE method is that because (5.3) is Gaussian, the estimated value of y_s is more probable to be the mean value of the function (Petrou and Sevilla, 2006). Therefore, from the linear least square sense the local residual will be,

$$\epsilon_s = y_s - \alpha_s \sum_{r\in N_s} y_{s+r} \tag{5.4}$$

For least square fitting, given a stationary texture, sample neighbourhoods of the texture are extracted by linear scanning of the estimation window w. Overlapping neighbourhoods are also allowed similar to the parameter estimation in LPH and RI-LPH feature extraction. Then the local least square solution is,

$$\hat{\alpha}_s = \arg\min_{\alpha_s} \sum_{t\in\Omega_s} \epsilon_t^2 = \arg\min_{\alpha_s} \sum_{t\in\Omega_s} \left(y_t - \alpha_s \sum_{r\in N} y_{t+r}\right)^2 \tag{5.5}$$

where Ω_s corresponds to the inside region of estimation window of size w. The regularization is not used here because the ill-conditioned problem due to the matrix non-invertibility does not occur here. By setting the first derivative to zero the local parameter value can be obtained as,

$$\hat{\alpha}_s = \frac{\sum_{t\in\Omega_s}\left(y_t \sum_{r\in N} y_{t+r}\right)}{\sum_{t\in\Omega_s}\left(\sum_{r\in N} y_{t+r}\right)^2} \tag{5.6}$$

The variance parameter of the model is then calculated by,

$$\hat{\sigma}_s^2 = \frac{1}{|\Omega_s|}\sum_{t\in\Omega_s}\left(y_t - \hat{\alpha}_s \sum_{r\in N} y_{t+r}\right)^2 \tag{5.7}$$

The simple forms of the solutions obtained for model parameters given by (5.6) and (5.7) can be easily implemented and efficiently computed. Estimating the parameter vectors $\mathbf{f}_s = [\hat{\alpha}_s, \hat{\sigma}_s]^T$, for all the pixels in Ω results in two parameter images corresponding to the two model parameters. The normalized histograms of each parameter image are then concatenated into one long vector which is used as the feature vector of the texture image.

5.3 Conclusions

Depending on the required output, some vision applications may need rotation invariant texture features. Rotation invariant features can be categorized as one of the omnidirectional features or the rotation invariant directional features. In this chapter methods of achieving rotation invariant features with the LPH descriptors are considered. Two descriptors, namely RI-LPH and I-LPH features are introduced which belong to the rotation invariant directional features and the omnidirectional features respectively. RI-LPH descriptors achieve rotational invariance via a circular shifted neighbour method and I-LPH features consider the IGMRF model to produce isotropic features. The performance of these features in texture segmentation and classification will be investigated in subsequent chapters.

Chapter 6

Texture Segmentation

This chapter considers the texture segmentation performance of the proposed texture descriptors. Texture segmentation involves dividing an image into different regions based on the texture characteristics. Texture discrimination as well as the region boundary localization are two main considerations in texture segmentation. This chapter deals with evaluating local parameter histogram descriptors in texture segmentation.

6.1 Default Variable Setting

The texture feature extraction process relies on few pre-defined variables which need to be adjusted manually. The values assigned for these variables are kept constant in all the experiments reported throughout this chapter unless stated otherwise.

For the formulation of LPH descriptors a neighbourhood size n should be defined. The neighbourhood size is kept as small as possible and set to $n = 3$ in this study. This is same as using a second order neighbourhood system with GMRFs which is commonly used in the GMRF literature. The estimation window size is selected as $w = 2n - 1$ to achieve consistency as discussed in Chapter 4. Thus, $w = 5$ is used when $n = 3$.

For RI-LPH and I-LPH descriptors, the radius $r = 1$ and the number of neighbours $p = 8$ is used. The estimation window size is calculated similar to the LPH descriptors. When r is given the value of n is $n = 2r + 1$. Therefore, the estimation window size in terms of r is $w = 4r + 1$ and here it is $w = 5$ when $r = 1$.

The local histogram construction window size b should be properly selected to minimize boundary localization errors while maximizing texture discrimination. Our experiments indicate that a window size in the range of $b = 9$ to $b = 23$ works well for fine to comparatively large texture patterns. Hence, $b = 11$ is selected for the following experiments. The number of bins for constructing a parameter histogram is fixed to $bins = 10$. The

bin images are smoothed using an average filter of size 7×7 as an additional post processing stage. The regularization parameter is set to $c = 1$. For extracting adaptive TGMRF features $n = 3$ and $w = 21$ is used following GMRF literature.

For semi-supervised segmentation, the k-means clustering algorithm is employed with a given number of clusters k. The k-means clustering algorithm is used here to emphasize the discriminative power of the texture descriptors alone, without inquiring much reinforcement from the segmentation method to the resulting segmented images. The L1-norm distance metric between histograms is used to measure the difference of two local parameter histograms. The k-means clustering algorithm disregards the spatial dependency property of adjacent neighbour labellings. Therefore, the spatial coordinates of the pixels are also used as two additional features which is a common technique used to overcome the negligence of spatial adjacency with k-means clustering based segmentation (Jain and Farrokhnia, 1990). The number of texture regions k is input by the user according to the number of texture types observed on the mosaic image.

The percentage error rate s_e is calculated as the ratio between the number of incorrectly segmented pixels and the total pixels in the image as given by the equation (6.1).

$$s_e = \frac{\text{no. of incorrectly segmented pixels}}{\text{total pixels in the image}} \times 100\% \tag{6.1}$$

6.2 General Texture Segmentation

6.2.1 Texture Datasets

For general texture segmentation mosaic datasets are employed to achieve quantitative evaluations of the proposed features. The mosaic images are constructed using four commonly used texture databases namely, Brodatz (Brodatz, 1966), Prague (Haindl and Mikeš, 2008), CUReT (gray) (Dana et al., 1997) and UIUC (Lazebnik et al., 2005). Grayscale mosaic images comprising four different texture regions are constructed. We refer to these datasets as BRODATZ, PRAGUE, CURET and UIUC respectively, throughout this chapter. Some sample mosaic images from each dataset are shown in Figure 6.1. BRODATZ, PRAGUE, CURET and UIUC mosaic datasets are comprised of 15, 60, 1380 and 240 mosaic images respectively. BRODATZ and CURET datasets contain fine to medium textures. PRAGUE and UIUC datasets contain fine to relatively large scale textures. CURET dataset contains many isotropic and similar looking texture classes.

6.2.2 Comparison to TGMRF Descriptors

One of the main objectives of this research is to demonstrate that by formulating local parameter distributions, improved descriptive texture descriptors compared to TGMRF

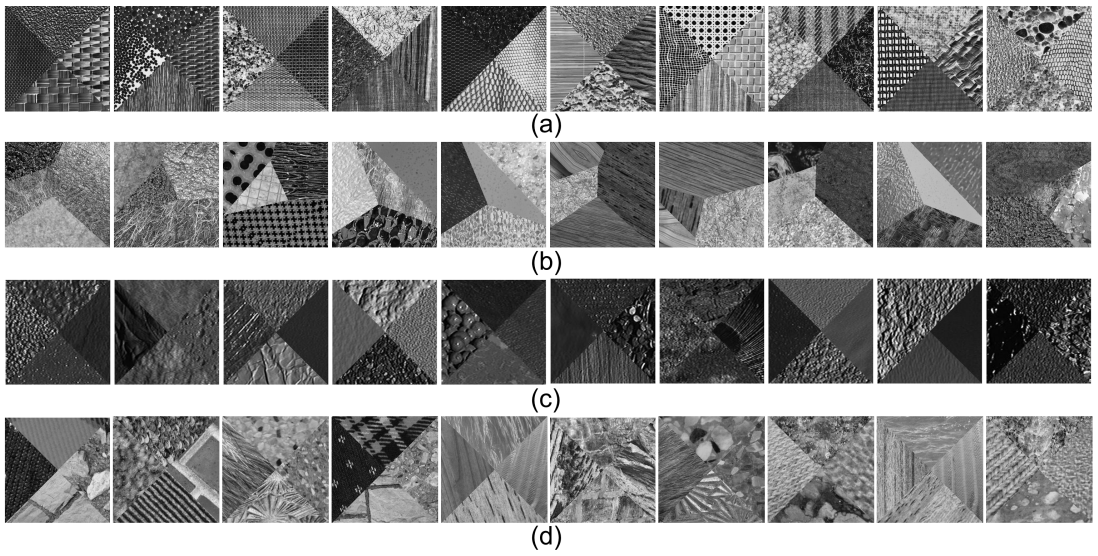


Figure 6.1: Sample mosaic images based on (a) BRODATZ, (b) PRAGUE, (c) CURET, and (d) UIUC textures.

Table 6.1: Segmentation error, $s_e(\%)$ for general texture segmentation.

Dataset		$s_e(\%)$			
		TGMRF	LPH	RI-LPH	I-LPH
BRODATZ	median	28.2	2.31	**1.66**	2.16
	IQR	9.01	0.92	1.08	1.85
CURET	median	27.9	4.33	**2.47**	2.55
	IQR	10.6	2.80	2.36	2.14
PRAGUE	median	32.6	**5.20**	7.55	10.9
	IQR	11.2	5.01	7.73	10.8
UIUC	median	36.0	7.84	**7.69**	10.3
	IQR	7.22	7.88	8.08	6.91

features can be constructed. Another objective is to examine the generalized performance of current TGMRF features on large datasets containing a variety of textures, specially medium to large scale texture patterns and deterministic textures.

First of all, the texture features are extracted from the mosaic images and then segmented using k-mean clustering algorithm. Next the s_e of each segmented mosaic image is calculated. The median of s_e and the interquartile range (IQR) are then calculated for each of the datasets. The results are shown in Table 6.1. The median is a more robust measure than the mean in this situation where texture segmentation is performed, because outlier values can effect the calculation of the mean. Outlier values which correspond to undesirably high s_e values may occur in the individual mosaic image segmentation tasks depending on the degree of difficulty of discriminating component textures in the mosaic image.

From Table 6.1, it is clearly observed that the local parameter distribution based features, LPH, RI-LPH, I-LPH descriptors outperform the discriminative power of adaptive TGMRF features. Among the local parameter distribution based features, the RI-LPH features perform slightly better than the other two descriptors. On average s_e of TGMRF descriptors is generally high and compared to BRODATZ and CURET datasets, the PRAGUE and UIUC datasets have a higher s_e. This may occur because the PRAGUE and UIUC datasets in general have relatively large texture patterns than the BRODATZ and CURET datasets. The small model order $n = 3$ used to construct the TGMRF descriptors may be inadequate to characterize the larger texture patterns. Also the PRAGUE dataset has challenging similar-looking component textures which could not be clearly differentiated by TGMRF descriptors which undergo an over-smoothed estimation process.

Considering the local parameter distribution based features, higher s_e are reported for PRAGUE and UIUC datasets compared to the BRODATZ and CURET datasets similar to the TGMRF descriptors (see Table 6.1). However, the s_e is significantly lower than the that obtained for TGMRF descriptors. The IQR values are also comparatively lower than that of TGMRF descriptors conveying more stable and robust segmentation results.

The PRAGUE and UIUC datasets in general have many directional texture patterns compared to BRODATZ and CURET datasets that are comprised of relatively fine and isotropic textures. Thus, I-LPH descriptors, which are suitable for more isotropic texture discrimination, perform well on BRODATZ and CURET datasets, however produce increased s_e for PRAGUE and UIUC datasets (see Table 6.1). A similar scenario can be observed on RI-LPH descriptors suggesting that the procedure of achieving rotation invariance may tend to affect negatively on the discriminative power of the features. Overall it seems that in texture segmentation, RI-LPH descriptors do not have a significant advantage over LPH descriptors as in a texture classification task which we will see in the next chapter. Therefore, features that are not rotation invariant may sometimes perform as good as its rotation invariant version in texture segmentation tasks.

6.2.3 Comparison to Other Texture Descriptors

In this section, we compare the segmentation performance of proposed features with other standard texture descriptors. The local binary patterns (LBP) is one of the popular state-of-the-art structural texture descriptors in texture analysis (Ojala et al., 2002). Also filter based Gabor texture descriptors are another well known method in texture analysis which closely relates to the biological vision system (Liu and Wang, 2003). These methods have been extensively analysed and used in many studies and applications in image processing and computer vision (Zhang et al., 2002; Sagiv et al., 2006; Liao and Chung, 2007; Sørensen et al., 2010; Zhu et al., 1998; Ahonen and Pietikäinen, 2008; Varma and Zisserman, 2009; Waller et al., 2011; Guo et al., 2010). Some studies have also

pointed out that these texture features can perform better than the TGMRF features (Ojala et al., 2001; Hadjidemetriou et al., 2003; Pietikäinen et al., 2000; Liu and Wang, 2003). Therefore, rotational invariant uniform local binary patterns (LBP) (Ojala et al., 2002) and spectral histograms (SH) (Liu and Wang, 2003) are employed in our study for the comparison. These features represent structural and spectral texture feature domains respectively and are also constructed based on local feature distributions.

For the filter based approach 52 filters are selected. Finding the correct filter set for the filter based approach is challenging. The filter set used here includes,

- 36 Gabor filters with six frequencies and six orientations
  - frequencies, $f = [0.14, 0.20, 0.23, 0.27, 0.30, 0.36]$ achieved according to frequency selection method proposed by Zhang et al. (2002)
  - orientations, $\theta = [0°, 30°, 60°, 90°, 120°, 150°]$
- 8 Gaussian filters
  - standard deviation, $\sigma = [\sqrt{2}/2, 1, 2, 3, 4, 5, 6, 10]$
- 8 Laplacian of Gaussian filters
  - standard deviation, $\sigma = [\sqrt{2}/2, 1, 2, 3, 4, 5, 6, 10]$ (Zhu et al., 1998)

The filter responses are contrast normalized according to Varma and Zisserman (2005). Each response image is converted to the corresponding spectral histogram features. i.e. the normalized histograms of the filter responses are used as the texture features (Liu and Wang, 2003). Features at each pixel are required for segmentation and are obtained by sliding window method. The histogram calculation window size of $b = 11$ is employed to calculate the pixel-wise features and coordinates are also used as two additional features, similar to the local parameter histogram based segmentation.

The rotational invariant uniform local binary patterns are calculated according to Ojala et al. (2002) and are implemented as in Heikkila and Ahonen (2012). The normalized histograms of local binary patterns are constructed using a histogram calculation window of size $b = 11$. Features from all three resolutions $(r, p) = (1, 8), (2, 16), (3, 24)$ are jointly used as final feature vector of a pixel. At each resolution, $bins = 10, 18, 26$ are used respectively as suggested in Ojala et al. (2002). The coordinates are also used as two additional features.

Our interest is on the texture features rather than the segmentation algorithm. Therefore, k-means clustering algorithm is kept constant as the default segmentation algorithm. However, other improved segmentation methods could further improve the segmentation results reported here. Also we have intentionally avoided using intensity

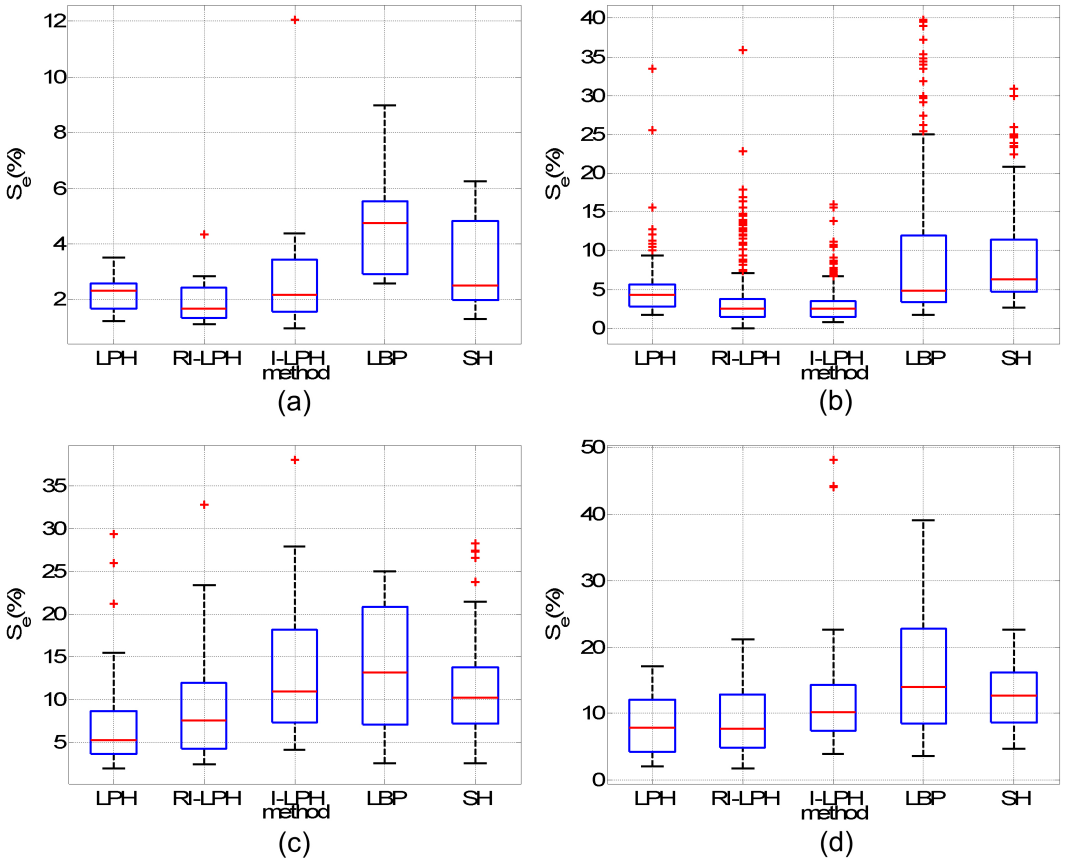


Figure 6.2: Segmentation error (s_e%) for mosaic datasets based on,(a) BRODATZ, (b) CURET, (c) PRAGUE and (d) UIUC textures.

information directly in any of the methods considered here and completely focused on texture based features for segmentation.

Figure 6.2 illustrates the box plots of s_e achieved for each dataset. In general we see from Figure 6.2 that LPH, RI-LPH and I-LPH descriptors have performed better compared to the other methods. BRODATZ and CURET datasets achieve lower s_e compared to the PRAGUE and UIUC datasets conveying that all descriptors better perform on more fine and more isotropic textures. The LBP descriptors achieve higher accuracies in classification (Ojala et al., 2002), however in these segmentation tasks their performance is lower than the other methods. The main reason for this may be that LBP features have less boundary localization quality than other features (Figure 6.3). Figure 6.3 shows sample segmented images from each of the methods. From Figure 6.3 it is observed that boundary localization of LBP with k-mean clustering algorithm is comparatively poor. Considering Figure 6.3e it is further noticed that LBP descriptors are less capable of capturing larger texture patterns. Also the construction of uniform LBP involves a small loss of discriminative information compared to rotation invariant LBP (Mäenpää and Pietikäinen, 2005). SH features perform better compared to LBP (see Figure 6.2), however larger filter sizes are required to capture larger patterns which in turn cause boundary localization errors (see Figure 6.3f). Nevertheless, SH features perform better

on larger textures as well as fine textures compared to LBP descriptors (Figure 6.2). Compared to LBP and SH descriptors, local parameter distribution based features perform better and are more capable of dealing with relatively larger texture patterns. Furthermore they have achieved a better boundary localization with the simple k-means clustering algorithm.

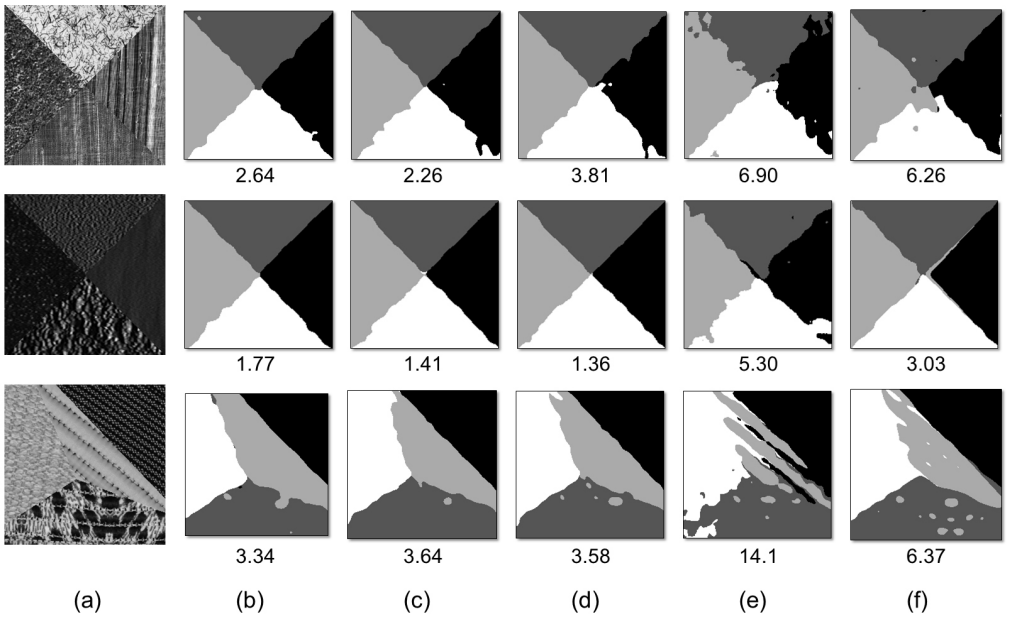


Figure 6.3: (a) Three randomly selected mosaic images. Segmented images for selected mosaic images using, (b) LPH, (c) RI-LPH, (d) I-LPH, (e) LBP and (f) SH texture descriptors. (s_e%) is given below each image.

In local parameter distribution based methods, a value for strength of local interactions are estimated. However, in LBP method a thresholding scheme is used to create binary signatures for neighbour interactions based on intensity differences. This reason could explain the better performance of local parameter based features over LBP features. Also construction of uniform binary patterns loose some texture information because of the restriction on the number of different patterns (Ojala et al., 2002).

For SH descriptors choosing an optimal filter set may lead to more accurate results. However, despite using a large filter set, SH descriptors have not performed better than LPH or RI-LPH descriptors. Therefore, local parameter distribution based features, specially LPH and RI-LPH features form more robust discriminative features compared to SH descriptors. Moreover, as seen in previous subsection LPH and RI-LPH descriptors have almost similar performances in texture segmentation and I-LPH features are more suitable for fine isotropic textures.

6.3 Variable Evaluation

Following subsections will examine the behaviour of local parameter distribution based features with different variables associated with their construction. A randomly selected subset of BRODATZ and CURET datasets are used for this purpose. All the other variables are kept in the default setting explained in section 6.1 except for the variable of interest.

6.3.1 Estimation Window Size

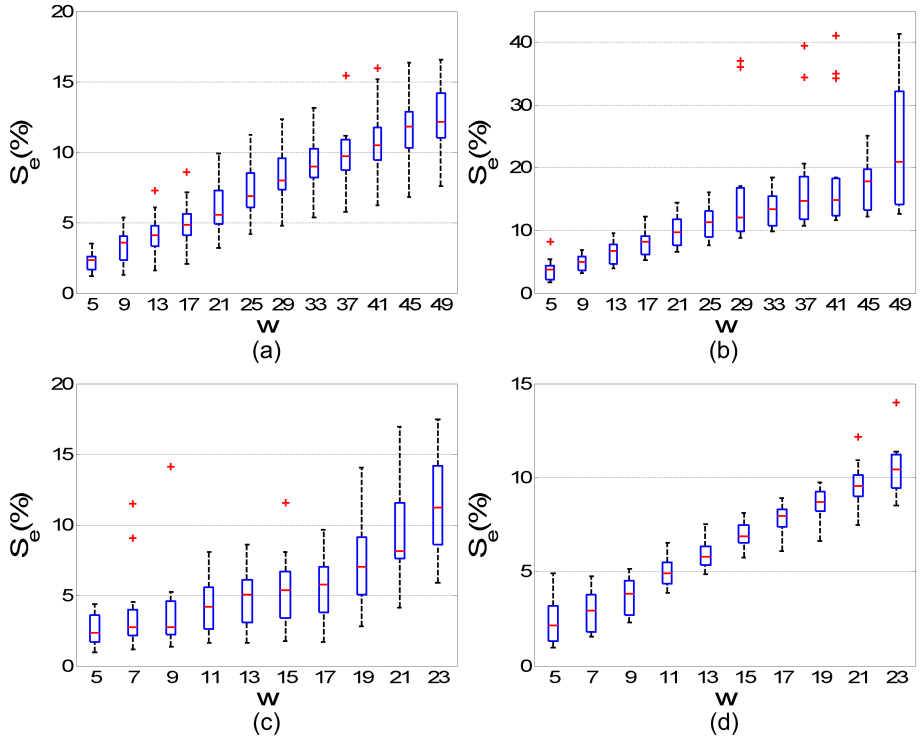


Figure 6.4: Segmentation error s_e with changing estimation window size w. (a) LPH on BRODATZ, (b) LPH on CURET, (c)I-LPH on BRODATZ, (d) I-LPH on CURET.

The estimation window size w is a critical variable in the construction process of local parameter distribution based features. The value of w should be small enough to capture structural details about the texture. When the value of w increases the estimates start to smooth out. Figure 6.4 illustrates how s_e increases with increasing estimation window size for LPH and I-LPH descriptors. The higher values of w here does not correspond to the TGMRF features. These are the segmentation results based on distributions of local parameters. $w = 21$ corresponds to the setting which TGMRF adaptive features are estimated. When using the TGMRF features it is seen that the s_e is around $25-30\%$ for BRODATZ and CURET datasets (see Table 6.1). But Figure 6.4 shows that constructing histograms of parameter estimates can reduce the s_e to roughly $5-15\%$.

Thus, local distribution construction contributes positively towards formulating a better texture feature. Additionally, smaller w values further improves the segmentation performances. Therefore, the rule introduced in Chapter 4 for selection of w to ensures a localized estimation process has further improved the segmentation results. The inconsistencies of using small estimation windows are alleviated by regularization. Hence, smaller w values improve the discriminative power of the local parameter distribution based features.

6.3.2 Regularization Parameter

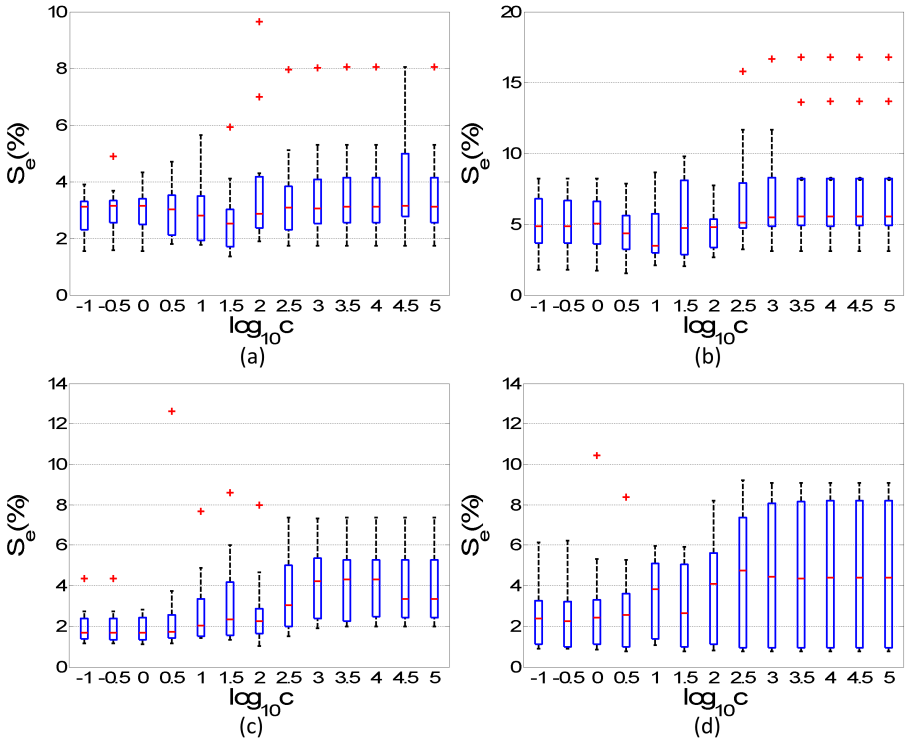


Figure 6.5: Segmentation error s_e with changing regularization parameter value c. (a) LPH on BRODATZ, (b) LPH on CURET, (c)RI-LPH on BRODATZ, (d) RI-LPH on CURET.

Regularization is critical in small model estimation where smaller estimation windows are employed leading to small sample sizes for the estimation process. Regularization is necessary to overcome the ill-posed problem associated with normal equations as discussed in Chapter 4. Thus, the main purpose of using regularization here is to assure the invertibility in least square problem rather than regularization of over-fitting problem. The default regularization parameter value is previously set to $c = 1$ heuristically based on prior knowledge of regularization techniques (Björkström, 2001).

Figure 6.5 demonstrates the effects of changing c on segmentation results achieved by LPH and RI-LPH descriptors. It is clear from Figure 6.5 that for a large range of c values segmentation error s_e remains fairly small. However, there exists a upper limit

for c value and according to Figure 6.5 it is in the range more than $c = 10^5$. Therefore, a value between 10^{-1} to 10^5 could be selected for c. Higher values of c reduce the estimated values of parameters towards zero. Thus, here $c = 1$ is an acceptable value for texture feature construction.

6.3.3 Histogram Calculation Window Size

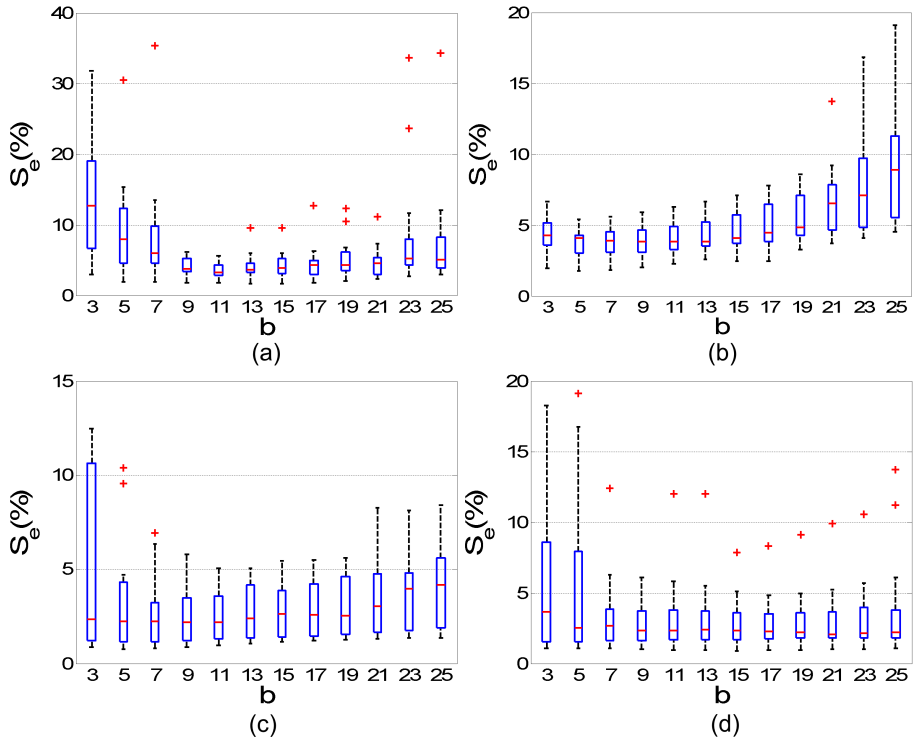


Figure 6.6: Segmentation error s_e with changing histogram calculation window size b. (a) LPH on BRODATZ, (b) LPH on CURET, (c)I-LPH on BRODATZ, (d) I-LPH on CURET.

Histogram calculation window size b defines the trade off between boundary localization and discriminative ability. Too small values for b could not capture the homogeneous characteristics of the texture and does not well represent the texture. Too high values for b on the other hand can capture the homogeneous texture pattern well, however increase the errors near the boundaries. Figure 6.6 shows the s_e with changing b for LPH and I-LPH features. It is clear from Figure 6.6 that a too small histogram calculation window size causes higher s_e because it cannot capture regional properties of the texture well. Nevertheless, when b is very large there is again a tendency to increase the s_e due to boundary localization errors. According to Figure 6.6 there exists a range of suitable b values. The b values between $b = 9$ to $b = 21$ is an acceptable range of b which properly compromises between discriminative ability and boundary localization problem. The selected default value, $b = 11$ for construction of local distribution based features is therefore a proper choice for histogram calculation window size.

6.4 Natural Image Segmentation

Natural images have textures that contain more noise and ambiguous variations. They also contain textured regions as well as smoothed and near piece-wise constant intensity regions. In human perception of natural scenes both colour and texture are efficiently processed to give an accurate understanding of the image regions. The dominating property of region recognition interchanges between colour and texture. For example, smoothed regions may be predominantly identified by their colour and regions with patterns are predominantly recognized by texture. Therefore, we integrate the colour information of the image to the texture features obtained from the gray-scale image. Colour and texture features are extracted separately.

The gray-scale images are used to extract the texture features. Employing texture features from the gray-scale version of the image is efficient and sufficient to capture the required properties of the texture. The default variable setting for LPH, RI-LPH and I-LPH descriptors are maintained. The color features are based on local averaged intensity histogram. First the color image is filtered using an averaged filter of size $w \times w$. This gives the averaged intensity at each pixel which also corresponds to the mean intensity value subtracted from the estimation window when estimating the small model parameters (recall that mean zero image region Ω_s is used to estimate parameters). On the averaged filtered image, by sliding a histogram calculation window of size $b \times b$ the local averaged intensity histograms are achieved. The values of w and b are same as the default values used in calculating the texture features.

When dealing with natural images it is observed that using $c = 1$ is not a suitable choice for controlling the regularization. Figure 6.7 illustrates the s_e with the regularization parameter c for two sample natural images and a general texture mosaic image. Because natural images have more variations in the textures and noisy components, the foreground and background regions have substantial intra-class variations. By controlling the regularization process we can penalize some degree of intra-class variations and still achieve fairly acceptable segmentations for the natural images. This is because the regularization process reduces the over-fitting quality of the small models and produces more generalized estimates for the small model parameters.

Considering Figure 6.7(b) it can be observed that for the two natural images, *img1* and *img2*, s_e is large when c is small suggesting that inhomogeneities in the textures interfere with the segmentation problem. However, when c is between $10^{\frac{3}{2}}$ to 10^4 the s_e is comparatively small implying intra-class variations have been smoothed out for some degree to achieve a much better segmentation. Nevertheless, beyond $c = 10^4$ the s_e again starts to increase conveying that estimates are over regularized and the texture descriptors no longer have discriminative capabilities. According to the degree of noise and distortion in the image the value of c can be controlled to obtained a satisfactory subjective segmented image. For general texture images such as *img3* in Figure 6.7,

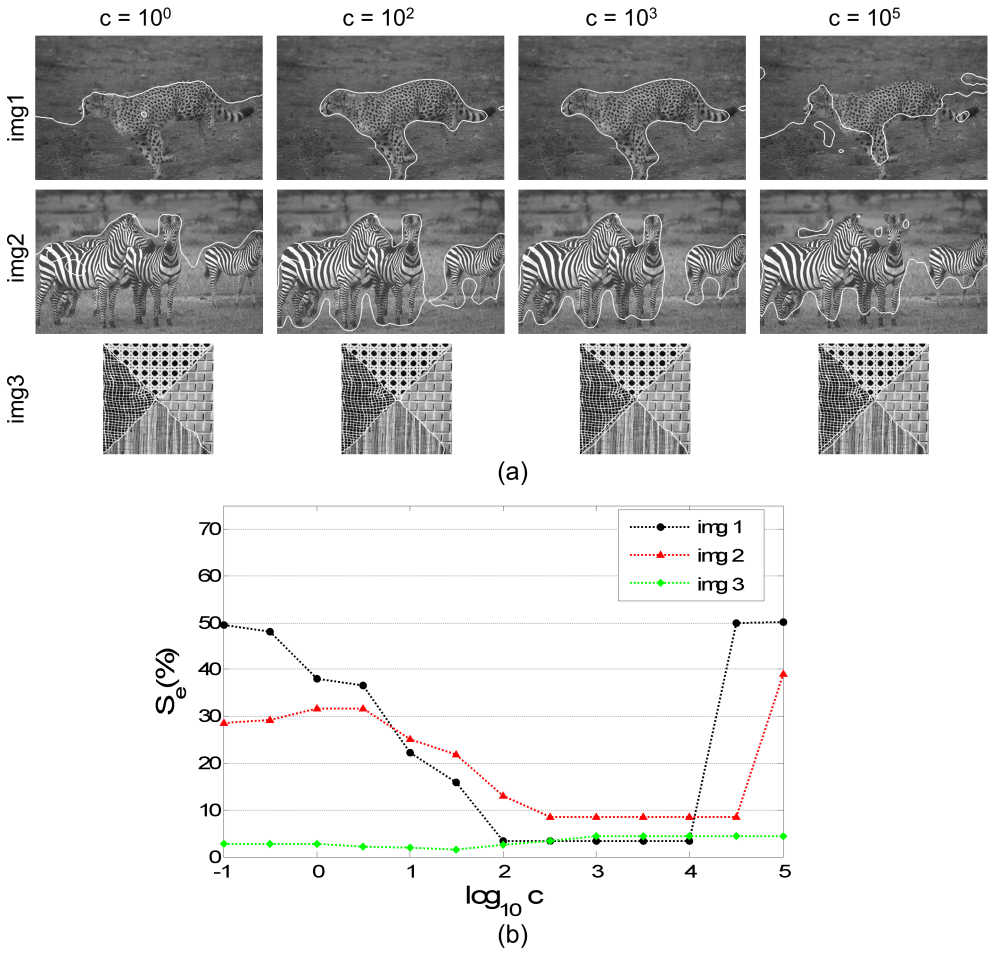


Figure 6.7: Effect of regularization parameter c. (a) Segmented images for different c values, (b) segmentation accuracy with c for three selected images.

the s_e is small throughout the range of c because the texture patterns are more regular without much disturbances from the outliers compared to that of natural images.

Thus, the regularization parameter value is increased to $c = 100$ for natural image segmentation. This setting performs well on many natural image segmentation problems where images are acquired under regular camera settings. The remaining variables are kept at the default setting (section 6.1). The results of colour image segmentation for sample images from Berkely dataset (Martin et al., 2001) are shown in Figure 6.8. The segmentations are achieved using LPH, RI-LPH, I-LPH and LBP descriptors. Note that the segmentation algorithm is the k-means clustering algorithm.

The segmentation results from LPH and RI-LPH descriptors with the averaged colour intensity histogram perform well compared to other descriptors (see Figure 6.8). I-LPH has perform well on more isotropic textures including smoothed intensity regions. Boundary localization is better in LPH and RI-LPH descriptors than that of I-LPH and LBP. LBP descriptors with the averaged colour intensity histogram have difficulties in discriminating more similar textures than the other methods. Here our objective

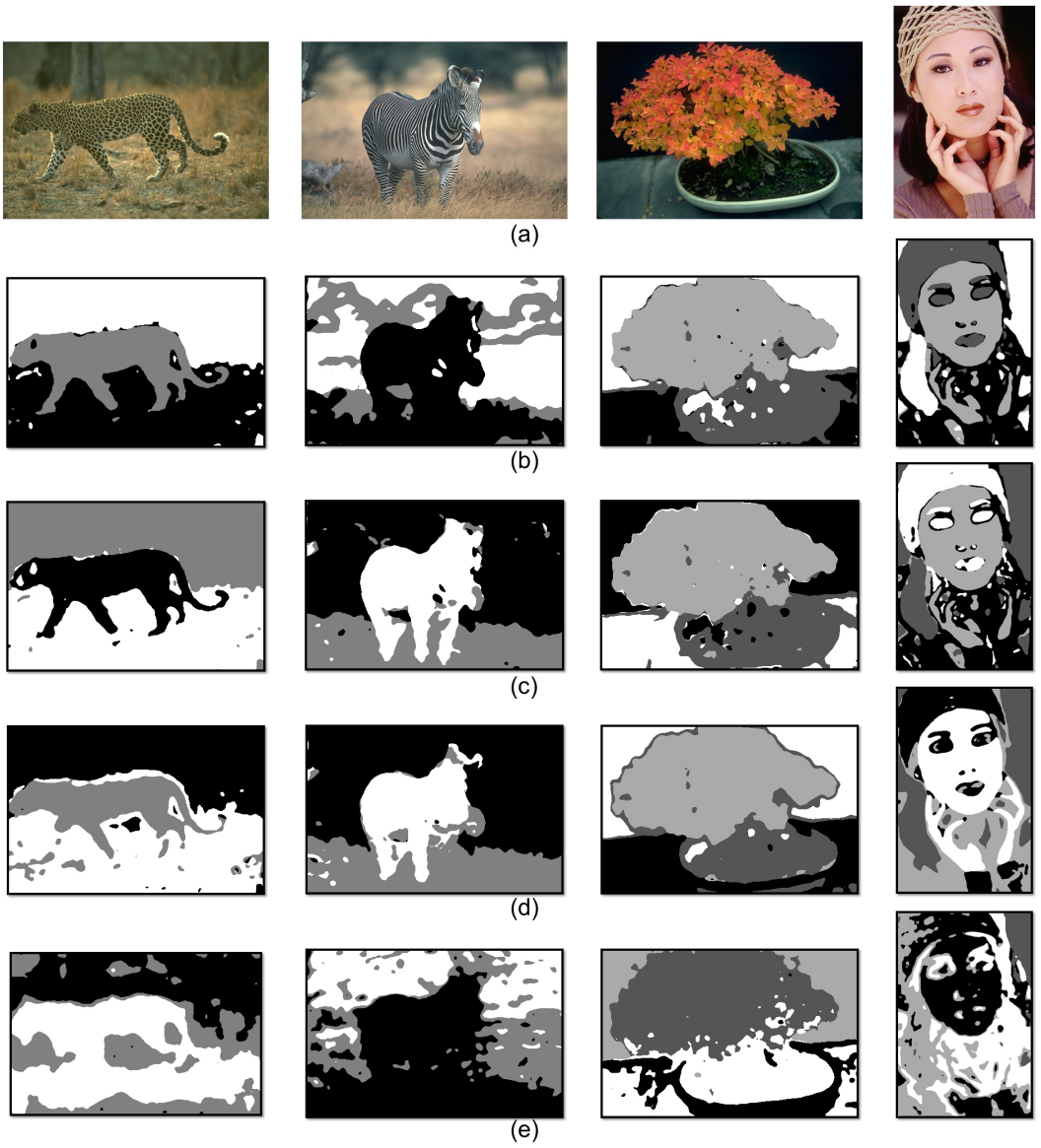


Figure 6.8: Natural image segmentation results achieved by different texture features. (a) Original image, (b) LPH (c) RI-LPH (d) I-LPH (e) LBP descriptors.

is to compare the performance of different descriptors under same setting rather than achieving perfect segmentations. However, using more advanced colour features and segmentation algorithms may further improve the results. More segmented images are shown in Figure 6.9.

6.4.1 Comparison to Other Methods

The semi-local region descriptor introduced by Houhou et al. (2009) and integrated active contours with Gabor features proposed by Sagiv et al. (2006) are used to compare the natural image segmentation performances. Both of these methods use gray-scale

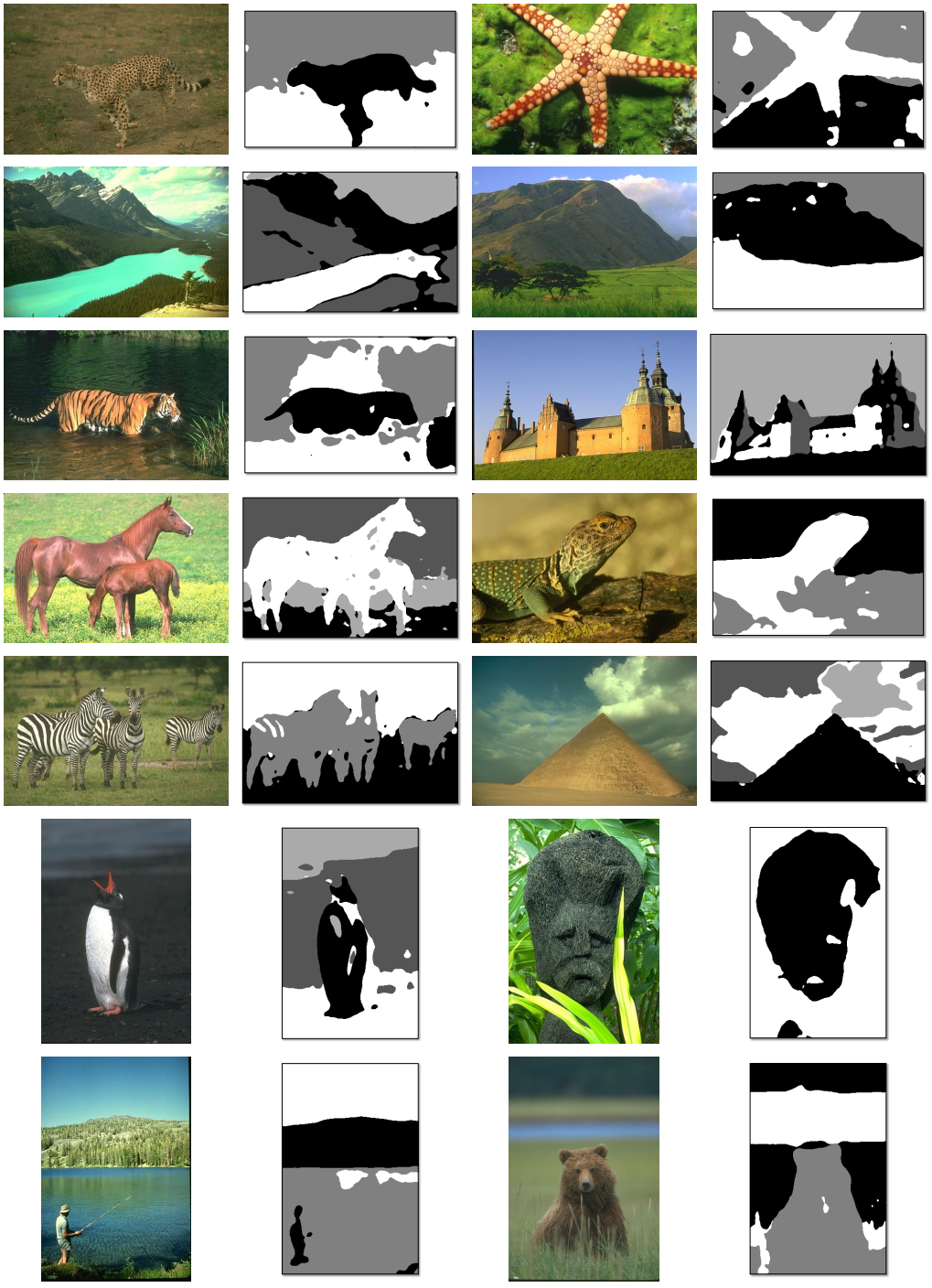


Figure 6.9: Colour image segmentation results obtained by using RI-LPH descriptors as texture features and k-mean clustering algorithm for segmentation.

natural images with the active contours as the segmentation algorithm. Therefore, we exclude the color features from the feature set and use the integrated active contours segmentation method proposed in Sagiv et al. (2006) for segmentation. Figure 6.10 shows some segmented results using RI-LPH descriptors with the results reported by semi-local region descriptors and Gabor feature based integrated active contours methods.

The results in Figure 6.10 shows that using RI-LPH descriptors are comparable with the results from semi-local region descriptors and in general better than the Gabor feature based integrated active contours. Notice that Gabor feature based integrated active contours and RI-LPH descriptors use the same segmentation algorithm, i.e. the integrated active contours (Appendix A).

6.5 Supervised Texture Segmentation

The segmentation tasks that has been carried out so far can be considered as semi-supervised segmentation problems. This is because they do not employ any training data to classify pixels in the image and use a clustering algorithm instead, which will divide the pixels into correct class based on the texture features. The segmentation process is not entirely unsupervised because the number of regions or the number of clusters need to be manually input by the user. There exist advanced segmentation algorithms which can automatically understand the number of segments and does not require user intervention to specify number of clusters. However, here our main concern is on the feature extraction method therefore such segmentation algorithms are not considered here.

It is interesting to carry out a supervised segmentation process where the user specifies the texture of interest and then the texture segmentation algorithm will segment out the regions in the image containing the texture of interest. Such a method have a practical value in image processing and worth investigating.

The process is supervised in the sense that a supervised training patch of interested texture, extracted from the input image itself, is fed into the system to calculate training data. The global parameters are fixed in the default setting for texture feature extraction.

The method is illustrated in Figure 6.11. First the supervised patch area is entered by the user by drawing a rectangular selection area which sufficiently captures the texture of interest. The coordinates of this area are stored by the algorithm for later use. Next the texture feature extraction is carried out for the whole image as previously discussed. RI-LPH descriptors are employed. Once the features are extracted for all pixels, the previously stored coordinates are used to identify the training feature vectors from the feature stack. Then the L1 norm distance between a feature vector of a pixel to its nearest training feature vector is calculated. The calculation of minimum distance to training models for all the image pixels results in the distance map. Figure 6.12b shows some distance maps obtained for some sample natural images. On the distance map lower values mean that the pixel is more probable to belong to the texture of interest.

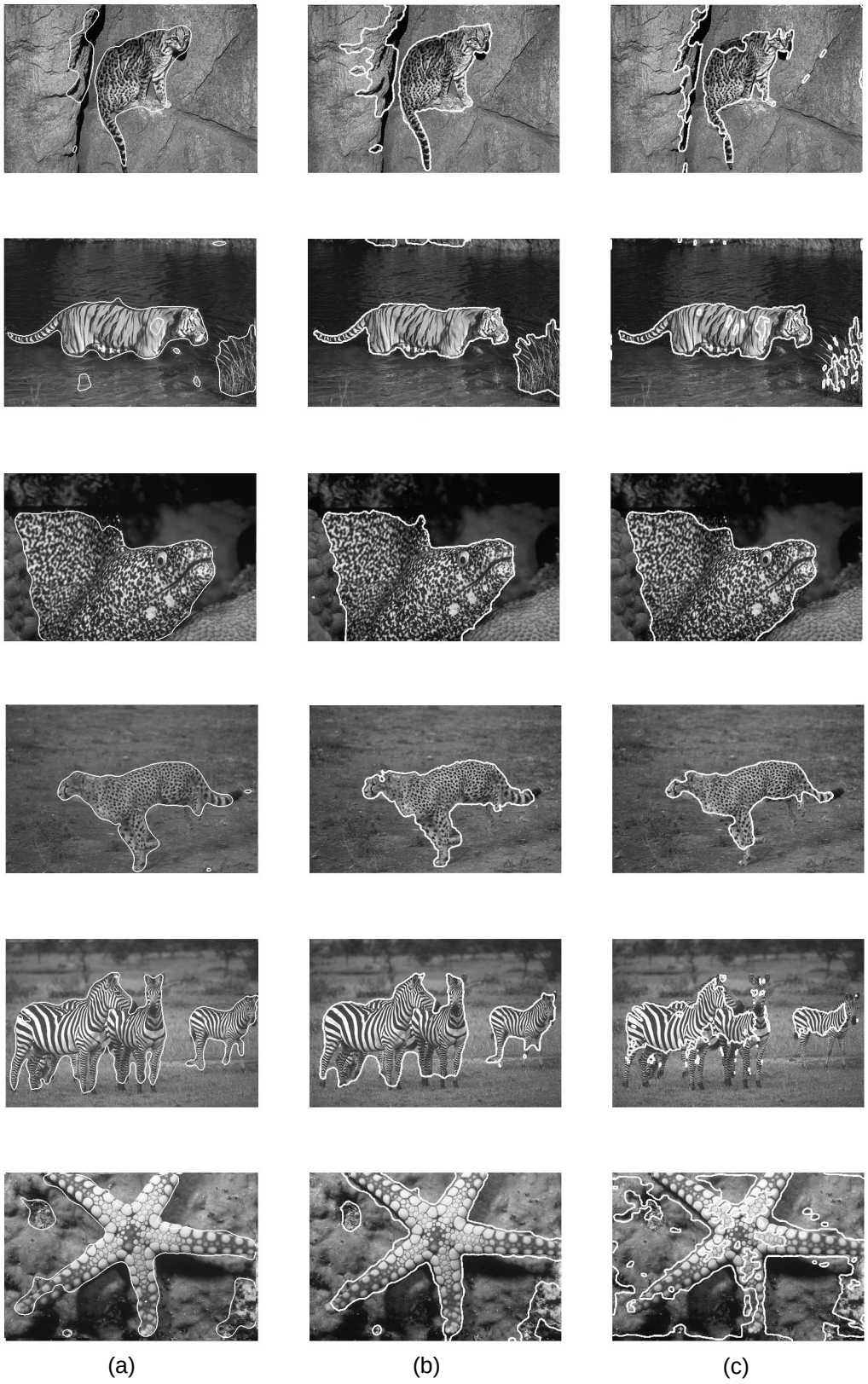


Figure 6.10: Gray scale natural image segmentation with active contour segmentation algorithm. (a) RI-LPH descriptors (b) Semi-Local Region Descriptors (Houhou et al., 2009) (b) Gabor Features (Sagiv et al., 2006).

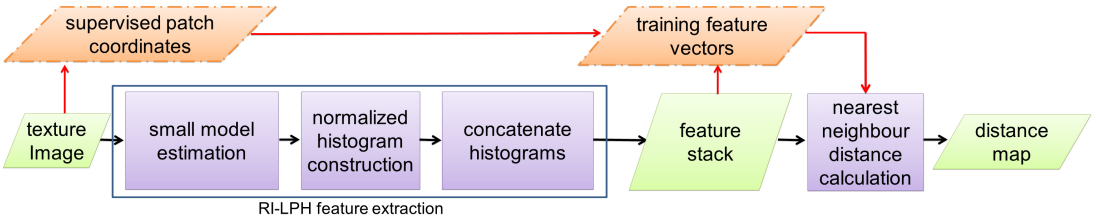


Figure 6.11: Supervised texture segmentation process.

After achieving the distance map, the integrated active contour segmentation algorithm is employed to find the boundaries of the region with the texture of interest (Sagiv et al., 2006). The segmentation results are shown in Figure 6.12c. Figure 6.12c demonstrates that the supervised segmentation method has been able to select out the regions with the texture of interest. Results are satisfactory, however it is seen that the final segmentation depends on how good the supervised training patch is. If the texture in the image have comparatively more variations which are not captured by the supervised training patch, sometimes the final segmentation have more errors. This problem can be alleviated by using more training patches possibly from the same image or other images if available. This technique is used in Chapter 8 in a medical problem and will be further examined. However, the proposed supervised segmentation method perform fairly well on segmenting out a region of specified texture.

6.6 Conclusions

The texture segmentation performance of improved texture descriptors proposed in this work is evaluated on commonly used general texture datasets and natural images. The local parameter distribution based features LPH, RI-LPH and I-LPH achieve significantly better performance compared to the TGMRF features. RI-LPH and LPH features have roughly similar performance in texture segmentation. Furthermore, I-LPH descriptors relatively perform well on more isotropic textures. Comparisons with LBP and SH descriptors, which are also based on local feature distributions, illustrate that LPH and RI-LPH descriptors form more robust discriminative texture features. Successful natural image segmentation can be also achieved using the proposed features. The segmentation error s_e with changing size of estimation window w, histogram calculation window b and the regularization parameter c are also investigated. The error s_e gradually increases with increasing w because the degree of parameter smoothing increases with the increasing w. On the other hand, there is a suitable range of values for b and c which have been used to select the default setting of the variables when constructing the features.

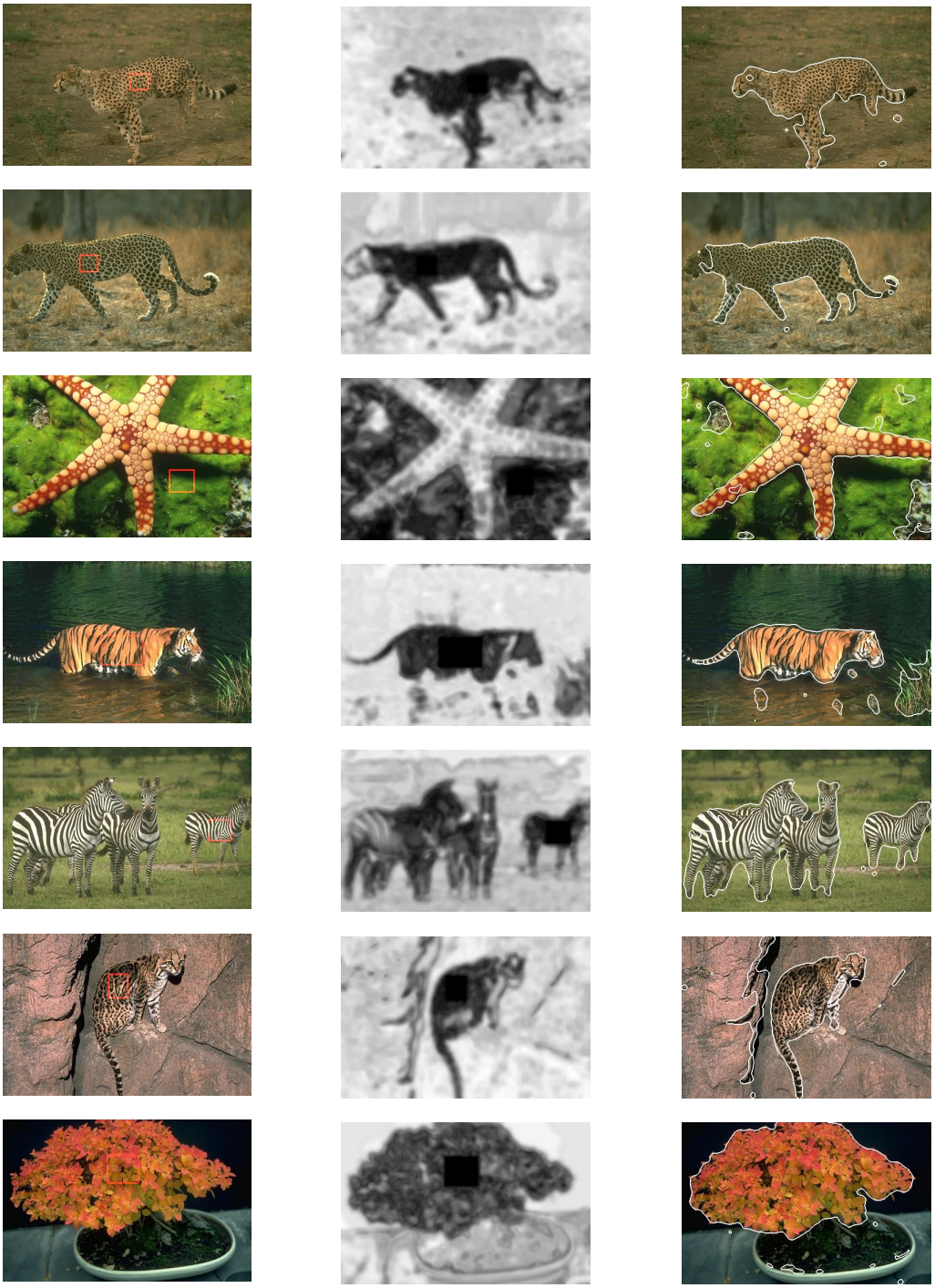


Figure 6.12: Supervised texture segmentation.(a) Original image (the selected area for training is marked by red square), (b) distance map, (c) final segmentation.

Chapter 7

Texture Classification

Texture classification categorizes an unknown texture image as belonging to one of a set of known classes depending on previously known training data. Classification can explain how well the features will perform in discriminating many texture classes, usually more than the number associated with a segmentation problem. This chapter focuses on classification performance of local parameter distribution based texture descriptors.

7.1 Default Variable Setting and Classifier

The default setting describes the pre-assigned variable values required for the feature extraction process and classification. If different values are used, it is explicitly mentioned in the content of the following subsections.

For construction of LPH descriptors $n = 7$ is used with $w = 2n - 1$ which is $w = 13$. For RI-LPH and I-LPH descriptors joint features from three resolutions $r = 1, 2, 3$ with number of neighbours in each resolution $p = 8$ is employed. The regularization parameter is kept as $c = 1$. For the histogram construction, the whole texture image is used which is equivalent to $b = imgSize$. For TGMRF descriptors, a neighbourhood size of $n = 7$ is employed. The estimation window size, when it is texture classification, usually takes the size of the image, therefore $w = imgSize$ is used.

In the classification experiments k-nearest neighbour (kNN) classifier with $k = 1$ is employed with the L1-norm distance metric. The kNN classifier is the simplest classifier possible and hence there is no special contribution from the classifier to the classification task. i.e., if the features give higher classification accuracies with kNN classifier, there is a higher chance of having even better results by more complex classifiers like support vector machines. The classification accuracies obtained here are for 100 classification problems repeated by choosing different training and test sets which are achieved by randomly dividing the dataset into equal class sizes (Varma and Zisserman, 2009). This

Table 7.1: Summary of texture datasets used for classification.

dataset name	no. of classes	images per class	total images	image size($pxls$)
BRODATZ	32	20	640	64×64
OUTEX	24	180	4320	128×128
CURET	61	92	5612	200×200

means that the training set includes rotation and scale variant instances of each texture class. The mean and standard deviation of accuracies over the 100 iterations are reported for conveying the statistical significance. Later on, invariant texture analysis is also considered where training set comprises of one rotated version of texture samples and the test set contains all texture samples from rest of the rotation angles.

7.2 Texture Datasets

Three commonly used datasets are employed for classification. A subset of Brodatz textures representing different homogeneous textures in the original Brodatz texture dataset was selected (Brodatz, 1966; Chen et al., 1995). This subset of Brodatz textures comprised of 32 texture classes similar to the dataset used in other studies including Valkealahti and Oja (1998) and Ojala et al. (2001). The full OUTEX_TC_00000 dataset having 24 OUTEX texture classes (Ojala et al., 2002) available from Outex Texture Database (2007) and full CUReT dataset (Dana et al., 1997) which comprises of 61 different texture classes are also used in the classification experiments. In this chapter, BRODATZ, OUTEX and CURET are used to refer to the datasets used in classification. BRODATZ data set was also used as the validation dataset for variable tuning.

The number of texture classes and samples associated with each dataset is shown in Table 7.1. All the images are histogram equalized before extracting the texture features as an additional pre-processing stage.

The 32 classes representing BRODATZ dataset are shown in Figure 7.1a and 12 sample images from two texture classes are shown in Figure 7.1b.

The OUTEX dataset contains textures from 9 different rotation angles and each angle contains 20 samples, therefore resulting 180 samples per class. The 24 OUTEX texture classes are shown in Figure 7.2a and 12 sample images from two texture classes are shown in Figure 7.2b.

The CURET dataset have rotation, scale and illumination variation in a class texture data. The 61 CURET texture classes are shown in Figure 7.3a and 12 sample images from two texture classes are shown in Figure 7.3b. Some texture classes are relatively

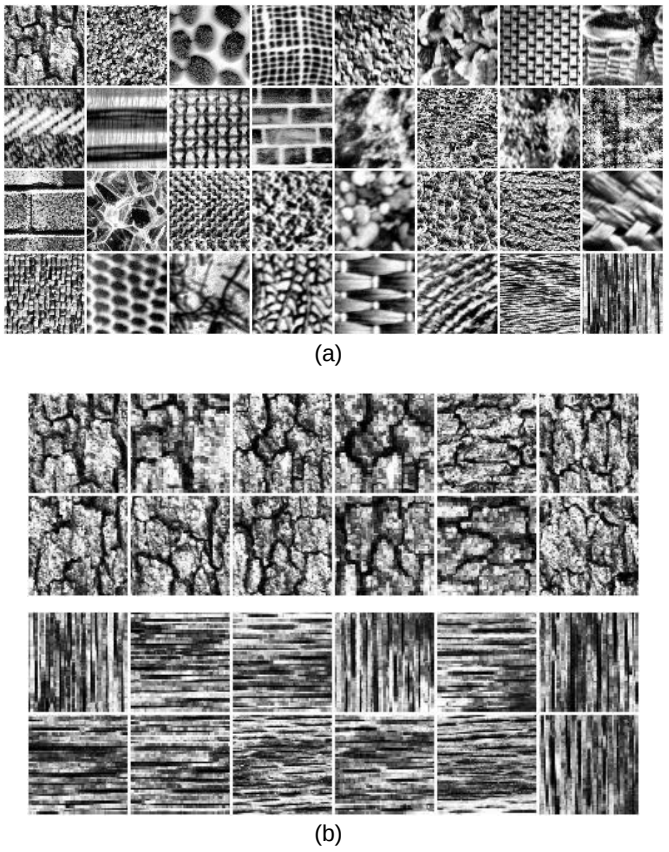


Figure 7.1: BRODATZ Dataset. (a) 32 texture classes, (b) samples from two random classes.

similar to each other making the classification task a challenging problem. Also CURET dataset has a large intra-class variation (Figure 7.3b).

These three datasets contain a large diversity of textures including directional, isotropic, large, fine, stochastic and deterministic characteristics. Thus, they are employed in our classification experiments to understand the general performance of the proposed texture descriptors. Only the texture cues are used in classification and direct use of intensity information are neglected.

7.3 Classification Results

7.3.1 Comparison to TGMRF Descriptors

First of all, the performance of GMRF based texture descriptors with neighbourhood size is examined. Figure 7.4 illustrates the classification accuracies for TGMRF and LPH descriptors with n and the classification accuracies for RI-LPH and I-LPH descriptors with r. From Figure 7.4 a,c and e it can be seen that the accuracy increases in general

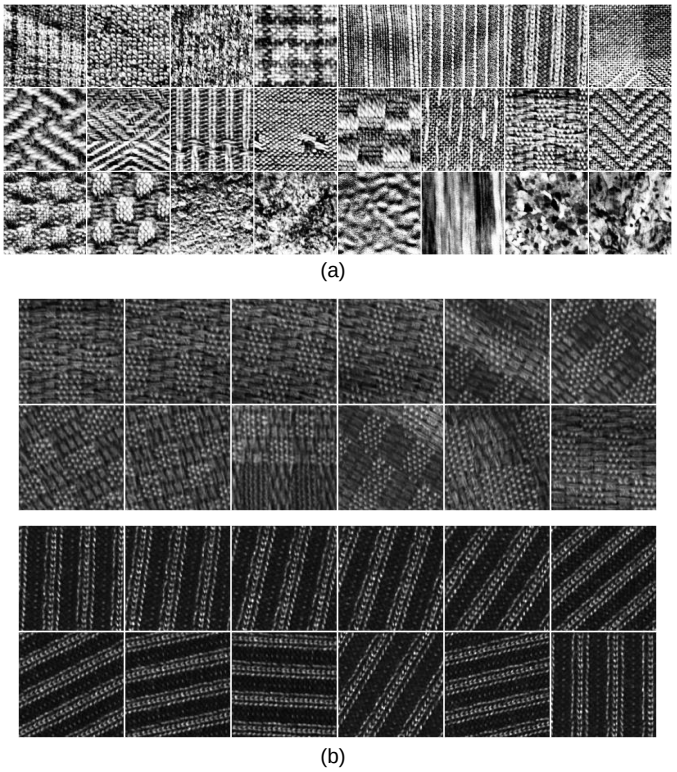


Figure 7.2: OUTEX Dataset. (a) 24 texture classes, (b) samples from two random classes.

for three datasets when neighbourhood size n increases for both TGMRF and LPH features. However, roughly after $n = 9$ the accuracy does not increase any more and it either saturates or declines. This conveys that low order features play a main role in texture characterization. Also LPH descriptors achieve a significant improvement over TGMRF texture descriptors.

From Figure 7.4 *b*,*d* and *f* on the other hand, a different behaviour can be observed for RI-LPH and I-LPH descriptors with r. In general for the three datasets, the accuracy slightly decreases or is the same with increasing r. Here the neighbourhoods represents a circular ring at r distance from the centre pixel and we assumed $p = 8$ neighbours in each level. Therefore, when the distance to the neighbours increases the accuracy decreases. The nearby neighbour pixels have a higher correlation with considered pixel relative to the far away neighbours. This implies that nearby neighbours are more important to formulate the feature descriptors which have stronger interactions with the considered centre pixel. However, the rate of accuracy reduction with r is small. Furthermore, Figure 7.4 illustrates that RI-LPH descriptors perform better than I-LPH descriptors.

Figure 7.4 demonstrates that when the features from different neighbourhoods are integrated, a higher classification accuracy can be achieved. Therefore, for RI-LPH and LPH descriptors $r = \{1, 2, 3\}$ feature sets are chosen as the texture descriptors. For TGMRF

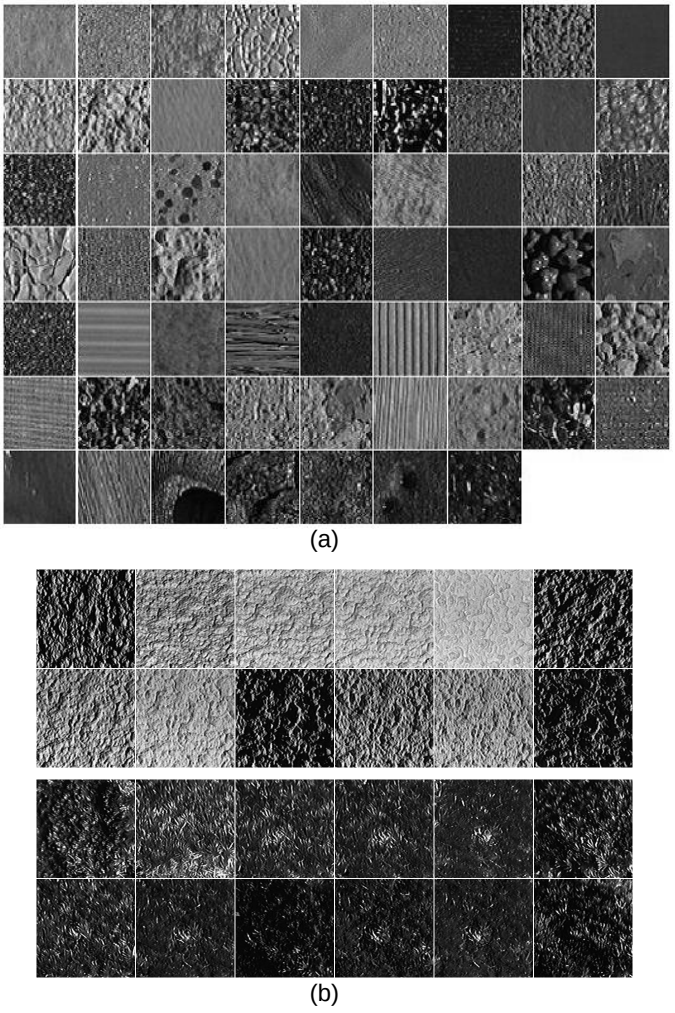


Figure 7.3: CURET Dataset. (a) 61 texture classes, (b) samples from two random classes.

and LPH however, we are limited to use $n = 7$ without integrating features from different neighbourhood sizes. This is mainly because when $n = 7$, it gives satisfactory levels of classification accuracies and it is also necessary to consider the computational cost associated with large neighbourhood sizes with many neighbours. Table 7.2 illustrates the classification accuracy of each descriptor under this setting which is also the default setting mentioned previously.

The results in the Table 7.2 are also graphically illustrated in Figure 7.5. It is seen from Figure 7.5 that RI-LPH descriptors perform better compared to other features because it is both rotation invariant and well suited for directional textures. Note that this experiment does not illustrate rotation invariant capabilities of RI-LPH descriptors properly, because we have included rotation variant instances in the training data. LPH and I-LPH descriptors have almost similar performances on each of the three datasets.

As expected, local parameter distribution based features LPH, RI-LPH and I-LPH descriptors have more discriminative ability than TGMRF texture descriptors. The MRFs

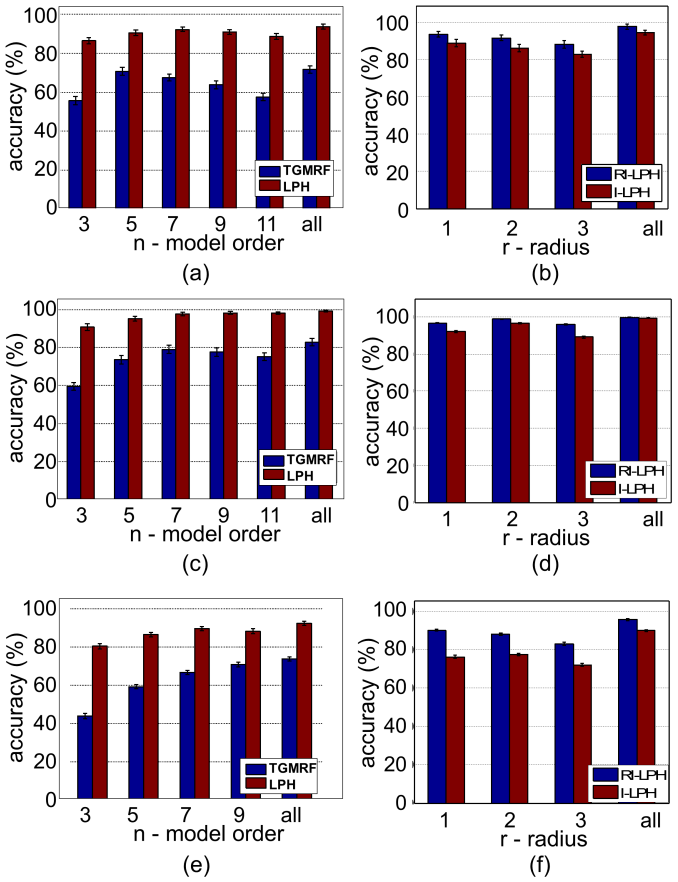


Figure 7.4: Texture classification accuracy with model order n and radius r. (a-b) BRODATZ, (c-d) OUTEX, (e-f) CURET dataset.

Table 7.2: Classification accuracies: Comparison with TGMRF descriptors. The mean classification accuracy and the standard deviation achieved from 100 repetitions of classification problem with equal size randomly divided training and test sets.

Dataset	method			
	TGMRF	LPH	RI-LPH	I-LPH
BRODATZ	68.1	92.4	98.0	94.4
	±2.10	±1.55	±1.03	±1.25
OUTEX	79.3	97.6	99.7	99.4
	±2.11	±0.87	±0.11	±0.18
CURET	67.4	89.1	95.6	89.4
	±0.76	±0.53	±0.36	±0.57

are powerful statistical tools of texture modelling and synthesis, however, a number of other studies have also reported their poor performance in texture segmentation and classification (Ojala et al., 2001; Hadjidemetriou et al., 2003). In texture classification and segmentation features should be more discriminative, and local parameter estimation and their distributions have lead to a solution to improve statistical model based

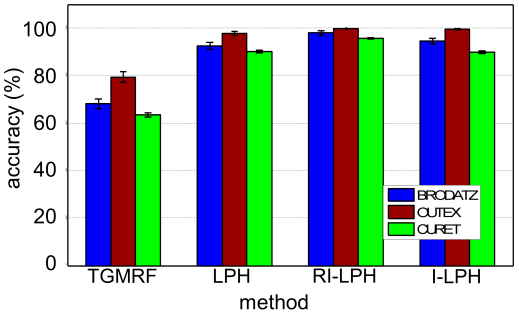


Figure 7.5: Texture classification accuracy: Comparison with TGMRF descriptors.

texture features.

7.3.2 Comparison to Other Texture Descriptors

The comparison of classification accuracies with accuracies from other texture feature extraction methods are considered here. The results are shown in Figure 7.6. The method SH uses the spectral histogram features and the method LBP employs the uniform local binary patterns as explained in Chapter 6. For construction of histograms the whole texture image is used. i.e., $b = imgSize$ similar to local parameter distribution based features. Direct intensity information is not employed here, allowing accuracy comparisons purely based on texture information. The classification accuracies are illustrated in Figure 7.6 for each dataset.

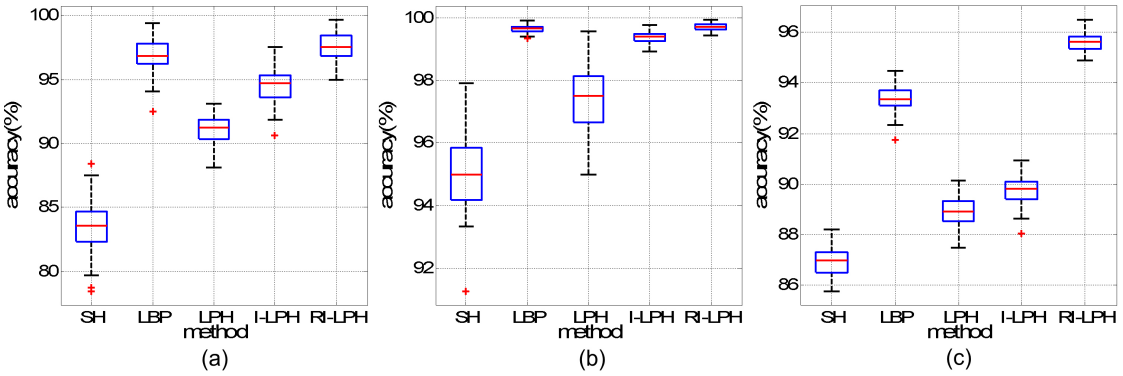


Figure 7.6: Texture classification accuracies: Comparison with other texture descriptors. (a) BRODATZ, (b) OUTEX, (c) CURET dataset.

Referring to Figure 7.6 it shows that RI-LPH features give the best classification accuracies for all datasets. I-LPH descriptors have performed comparatively lower than RI-LPH, because I-LPH descriptors ignores directional information and is more suitable for isotropic textures. For SH descriptors a reduction of accuracy is observed. This may be the case that the filter set we have chosen is not optimal, despite the fact that it is a fairly large filter set. However, optimal filter selection is a challenging process.

Table 7.3: Classification accuracies (%) reported in the literature for a verity of existing texture descriptors.

method	dataset		
	BRODATZ	OUTEX	CURET
RI-LPH[1]	98.1	99.7	97.0
TGMRF (7^{th} order) (Ojala et al., 2001)	71.3		
LBP (Ojala et al., 2002)		99.5	
VZ-joint (Varma and Zisserman, 2009)			97.2
VZ-MR8 (Varma and Zisserman, 2005)			96.3
signed difference (Ojala et al., 2001)	96.8		
derivative filters (Tou et al., 2009)	84.7		
GLCM (Tou et al., 2009)	85.7		
higher order GLCM (Ojala et al., 2001)	94.4		
Gabor features (Patil et al., 2013)	91.8		
Gabor features (Ojala et al., 2001)	93.6		
completed LBP (Guo and Zhang, 2010)		99.1	97.4

[1] integrating average intensity histogram as discussed in Chapter 6.

Also we have intentionally avoided using the intensity histogram, which may reduce the performance of SH descriptors. Nevertheless, LBP descriptors perform well for all the three datasets. Further comparisons of the texture descriptors including other texture descriptors are shown in Table 7.3. Clearly TGMRF descriptors have a poor generalized performance (Table 7.3), however through this research we have been able to formulate improved GMRF based texture features which produce comparable results with existing texture descriptors.

7.3.3 Rotation Invariant Analysis

The performance of rotation invariant texture classification with the proposed texture features is considered next. This experiment can determine whether the extension of LPH descriptors to the rotation invariance features through RI-LPH and I-LPH descriptors actually achieve robust performance against rotation variation. We use the OUTEX dataset for this task, which has textures from nine different angles namely, $0, 5, 10, 15, 30, 45, 60, 75$ and 90 degrees. Each angle has 20 sample images, hence 180 ($= 9 \times 20$) samples per class (see table 7.1). The texture samples from one particular angle is used for training and rest of the samples are used for testing. i.e. 20 samples from each class are used in the training set and 160 samples from each class are used in the testing set. The classification accuracies are shown in Figure 7.7.

It can be seen from Figure 7.7 that RI-LPH and I-LPH descriptors achieve reasonably good rotation invariance giving better classification accuracies compared to original LPH descriptors. The results in Figure 7.7 conveys that the training data from one rotation angle of the texture can successfully classify the rest of the dataset correctly by using

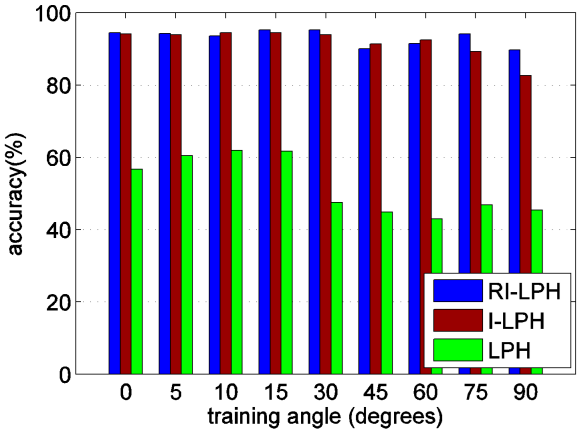


Figure 7.7: Classification accuracy with the training angle θ.

Table 7.4: Classification accuracies achieved with different rotation angles as the training set (maximum and minimum values are in bold font).

training angle	descriptor		
	LPH	RI-LPH	I-LPH
0°	56.7	94.4	94.2
5°	60.5	94.2	93.9
10°	**61.9**	93.5	94.5
15°	61.7	**95.3**	**94.6**
30°	47.5	95.2	93.9
45°	44.8	90.1	91.4
60°	**43.0**	91.5	92.5
75°	46.8	94.1	89.2
90°	45.4	**89.7**	**82.7**

RI-LPH and I-LPH descriptors. The classification accuracies are also shown in Table 7.4. In general, with the LPH descriptors which are not rotation invariant, the accuracy remains between $40 - 60\%$ across all the tests with different rotation angles for the training set. However, with RI-LPH the accuracy maintains between $90 - 95\%$ and when I-LPH is used the accuracy stays between $80 - 95\%$. This means that the RI-LPH and I-LPH descriptors can obtain good classification performances when the classifier is trained using one rotation angle of texture and tested on samples from different rotaion angles other than the one selected to train. Therefore, RI-LPH and I-LPH descriptors can be employed as rotation invariant texture descriptors.

7.3.4 Time Consumption

In order to analyse the efficiency of each feature extraction method, the time consumption for feature extraction is examined. Here, the time elapsed to extract texture features from a texture image of size 200×200 in a Matlab R2013a environment running on a

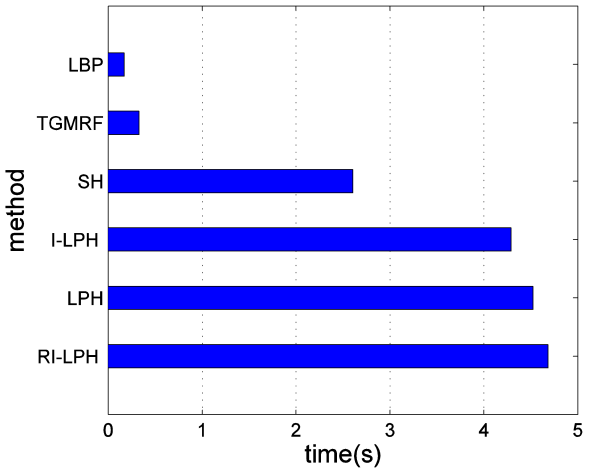


Figure 7.8: Time elapsed to extract texture features from a image of size 200 × 200 using different texture feature extraction methods.

2.67 GHz CPU with 12GB RAM is reported. Figure 7.8 highlights the time consumption comparison. From the Figure 7.8 it is observed that LPH, RI-LPH and I-LPH descriptors are computationally expensive compared to the other features. This is one of the weaknesses of local parameter distribution based features compared to TGMRF descriptors. Because local parameter distribution based features involve local parameter estimation at each pixel and additionally have a histogram construction stage, the computational cost is relatively higher. Nevertheless, the difference between the computation times of TGMRF features and LPH, RI-LPH and I-LPH descriptors are a few seconds. Therefore, the computation of LPH, RI-LPH and I-LPH descriptors are still practically reasonable.

7.3.5 Estimation Window Size

Next we look at the classification performance with changing size of the estimation window, w which is one of the critical variables in estimating small models. Figure 7.9 illustrates the classification accuracy with w for the three datasets BRODATZ, OUTEX and CURET. Here, RI-LPH and I-LPH descriptors are considered with $(r, p) = (1, 8)$ setting. The RI-LPH and I-LPH descriptors have better performance when the estimation window size is smallest possible (Figure 7.9) similar to the results from segmentation in Chapter 6. When the estimation window is small the smoothing of local structural features is minimum and therefore can achieve better performance with local parameter distribution based features.

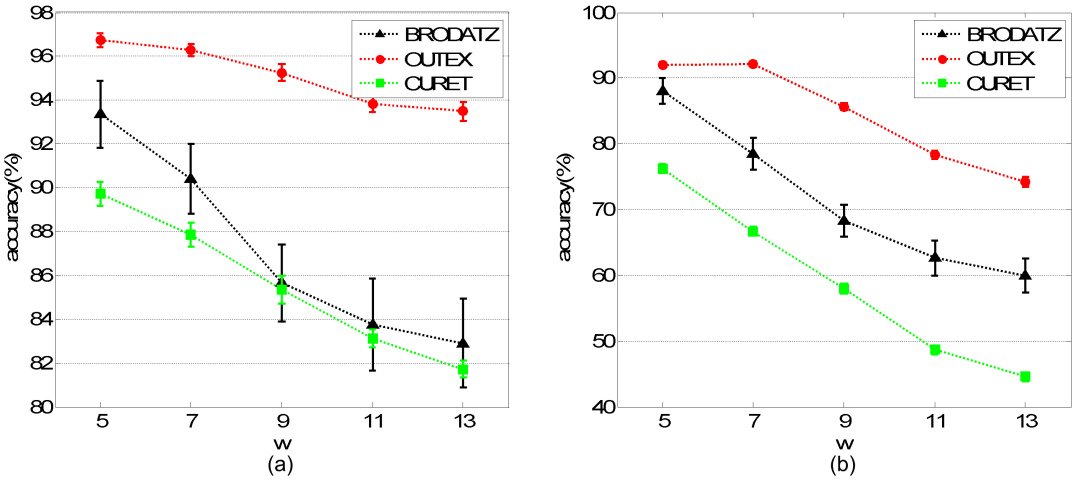


Figure 7.9: Classification accuracy with estimation window size w. (a) RI-LPH, (b) I-LPH. (for $r = 1, p = 8$ resolution).

7.3.6 Regularization Parameter

We have set the value of the regularization parameter manually as $c = 1$ in all the above experiments. However, here we analyse the classification accuracy with c. Figure 7.10 shows the variation of accuracy with c for the three datasets achieved using RI-LPH descriptors. When c is very large, it over smooths the parameter estimation process by reducing the value of the estimates towards zero. Therefore, the discriminative ability of the descriptors are slightly reduced. From Figure 7.10 it can be concluded that $c = 1$ is a proper choice for the regularization parameter for general texture classification with RI-LPH descriptors similar to general texture segmentation.

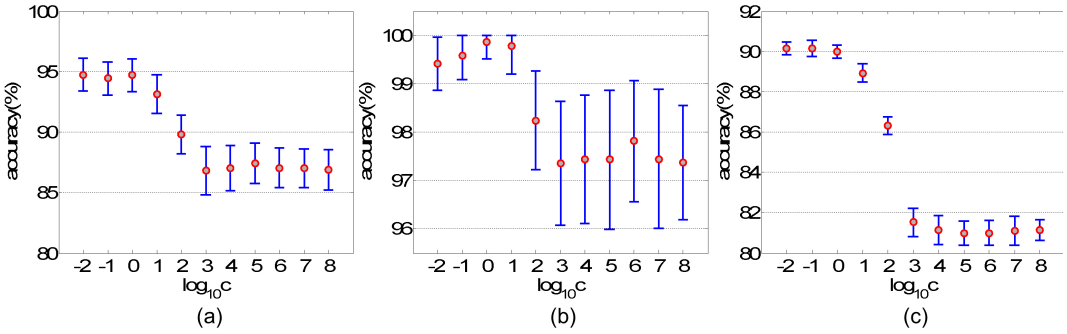


Figure 7.10: Classification accuracy with the regularization parameter c for RI-LPH descriptors (for $r = 1, p = 8$ resolution). (a) BRODATZ, (b) OUTEX, (c) CURET dataset.

7.4 Conclusions

Texture classification can explain how well the features will perform in discriminating many texture classes, usually more than the number associated with texture segmentation. We have analysed the classification performance of LPH, I-LPH and RI-LPH features using popular texture datasets. The classification results demonstrate that the RI-LPH features achieve slightly higher accuracies compared to LPH and I-LPH descriptors. Also these accuracies are significantly higher than that of TGMRF features. Furthermore, RI-LPH descriptors achieve better classification rates compared to SH features and comparable results with LBP descriptors. Comparison with other methods shows that RI-LPH features can achieve comparable classification rates with the state-of-the-art performances. Rotation invariant classification results illustrate that RI-LPH and I-LPH features can be employed as rotation invariant texture descriptors.

Chapter 8

Application to Emphysema Quantification

In this chapter we apply the local parameter distribution based descriptors on a real world application. Medical image processing is a key area of study where texture analysis is highly valued. Here we look at how to perform lung tissue quantification to diagnose a specific lung disease, given High Resolution Computed Tomography (HRCT) data using the texture feature extraction method proposed in this study. Emphysema is a critical lung disease causing extensive lung tissue destruction and is currently emerging as a worldwide health problem. Here, the I-LPH descriptors are selected as the preferred choice of texture descriptors because the parameter estimation of IGMRFs is simple and fast compared to other MRFs and ideally suited for isotropic texture representation which is the case with the considered medical condition. The approach is used to classify between healthy lung tissue (NT) and two sub types of emphysema, namely centrilobular emphysema and paraseptal emphysema. Local parameter distributions are more suitable in this application because it requires capturing subtle pattern variations which are crucial in tissue discrimination which may smoothed out by TGMRF descriptors. We also focus on supervised lung tissue segmentation to achieve lung parenchyma pathology distributions and quantification of emphysema and its subtypes.

8.1 Emphysema

Emphysema is a common lung disease which fatally disturbs the respiratory process due to permanent destruction of the lung tissue. Emphysema belongs to a family of lung diseases called Chronic Obstructive Pulmonary Diseases (COPD) which is defined as slowly progressive obstructive lung diseases in which the air exchange is impaired by the narrowing of the lower airways or destruction of alveoli.

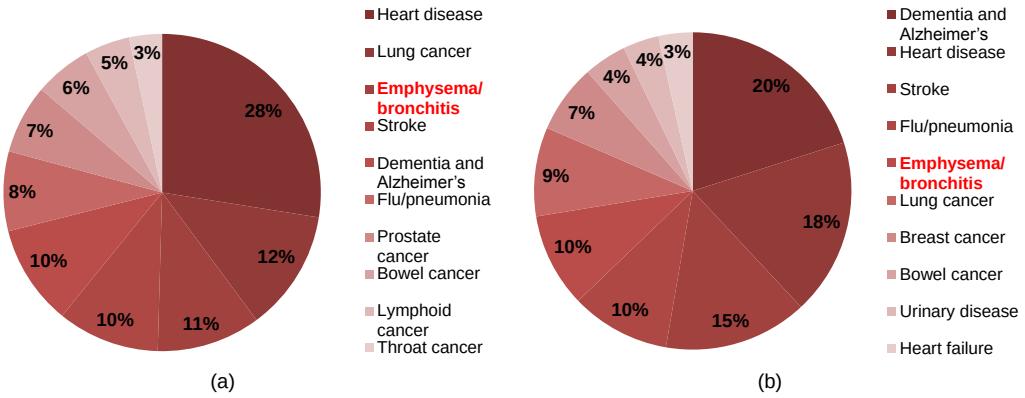


Figure 8.1: Leading causes of death in England and Wales, 2012. (a) mortality rate for men, (b) mortality rate for women, in percentages. Produced using data from Office for National Statistics (2013).

The main symptoms include shortness of breath, coughing, wheezing, weight loss and the tightness in the chest. If left untreated it may cause respiratory failure or heart failure. Typically the shortness of breath is worse on exhalation with a notable prolonged duration and worsens over time. A barrel chest is a characteristic sign of emphysema, but is relatively uncommon (Wikipedia, 2014; Duck, 2008).

Worldwide, COPD affects 329 million people or nearly 5% of the population. In 2011, it ranked as the fourth-leading cause of death, killing over 3 million people (Wikipedia, 2014). The number of deaths is expected to increase due to higher smoking rates and the aging population in many countries. The leading causes of death reported in England and Wales in the year 2012 is shown in Figure 8.1. This gives a rough idea about the serious impact of emphysema as a major health issue.

8.1.1 Causes of Emphysema

The main cause of emphysema is tobacco smoke. Of those with emphysema, 80% are either current smokers or have previously smoked (Data fact sheet, 2003). The likelihood of developing COPD increases with the total smoke exposure. In non-smokers, secondhand smoking increases the risk of having emphysema.

Poorly ventilated cooking fires, often fuelled by coal or biomass fuels such as wood and animal dung, lead to indoor air pollution and are one of the most common causes of COPD in developing countries. Furthermore, people who live in large cities have a higher rate of emphysema compared to people who live in rural areas (Wikipedia, 2014).

Intense and prolonged exposure to workplace dusts, chemicals and fumes increase the risk of emphysema in both smokers and non smokers. In some professions the risks have been estimated as equivalent to that of half to two packs of cigarettes a day.

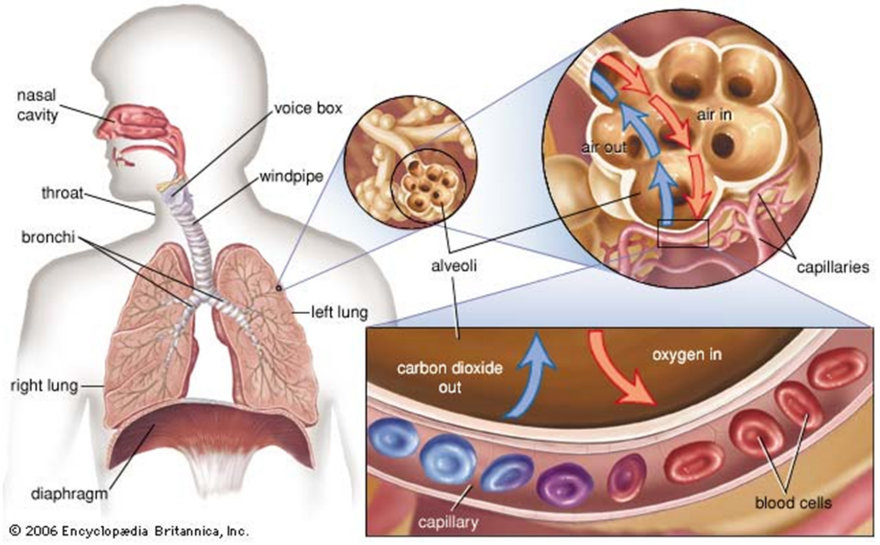


Figure 8.2: Structure of the human lungs (By courtesy of Encyclopaedia Britannica, Inc., copyright 2006; used with permission.) (Encyclopaedia Britannica, 2006).

A genetic disorder called Alpa-1-Antitrypsin deficiency (AAT) also leads to emphysema, but this is less common. The risk is particularly high if someone deficient in alpha 1-antitrypsin also smokes. In the United States, AAT deficiency is the primary cause of only 1% to 2% of cases of COPD (Duck, 2008).

8.1.2 Pathophysiology

The American Thoracic Society has define emphysema as 'the permanent enlargement of the air spaces, distal to the terminal bronchiole as a result of destruction of alveolar walls without significant fibrosis' (Lynch et al., 2000). Emphysema is characterized by destruction of alveoli, its walls and elastic fibre. A brief explanation about pathophysiology of emphysema is given in this subsection (Hasudungan, 2014).

The respiratory system consists of anatomical structures involved in ventilation and gas exchange. It allows to breath oxygen and exhale carbon dioxide. The gas exchange occurs through the alveoli, the tiny sacks-like structures distal to terminal bronchioles (Figure 8.2).

The toxins in cigarette smoke or polluted air cause an immune response process in the alveoli which releases chemicals that cause destruction of alveoli walls and the elastic fibres (Figure 8.3). The destruction of alveoli walls reduce the effective area in the lung for gas exchange. The elastic fibres allow recoiling during exhalation and inhalation. The destruction of elastic fibre causes narrowing of bronchioles and reduced recoiling of the alveoli which leads to trapped air inside the alveoli. The amount of trapped air inside the alveoli increases with time and results in inflated alveoli. These changes cause

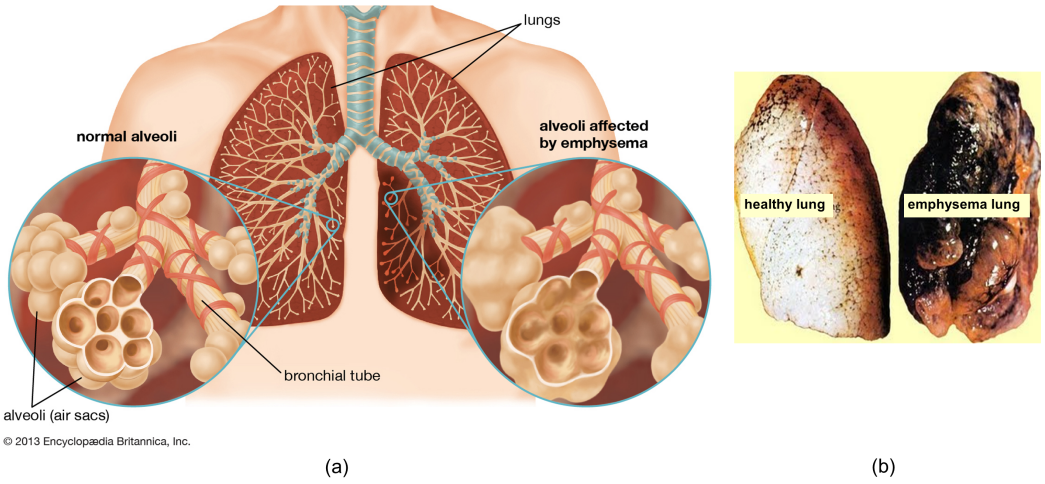


Figure 8.3: (a) Illustration of normal alveoli and alveoli affected by emphysema (By courtesy of Encyclopaedia Britannica, Inc., copyright 2013; used with permission) (Encyclopaedia Britannica, 2013), (b) images of the normal lung and the emphysema lung.

difficulty in breathing and in advance stages of emphysema it may cause hypoxia due to inadequate oxygen supply.

8.2 Diagnostic Methods

There are mainly two diagnostic approaches for emphysema, namely Pulmonary Function Test (PFT) also called spirometry and Computed Tomography (CT) image analysis. Following subsections will discuss these methods in details.

8.2.1 Pulmonary Function Test

Spirometry is the first and most common clinical lung function test to diagnose emphysema. It measures the volume of air and the air flow in exhalation. During the test, the subject will breathe into a mouthpiece attached to a recording device called the spirometer.

The more common lung function values measured with spirometry are Forced Vital Capacity (FVC) and Forced Expiratory Volume (FEV). The FVC measures the amount of air the subject exhale with force after inhaling as deeply as possible. The FEV measures the amount of air the subject can exhale with force in one breath. The amount of air the subject exhale may be measured at 1 second (FEV_1), 2 seconds (FEV_2), or 3 seconds (FEV_3). The severity of emphysema is achieved based on these measurements, specially FEV_1. Table 8.1 shows the severity stages of emphysema and the corresponding FEV_1 limits.

Table 8.1: Severity stages of emphysema. (GOLD-Global Initiative on Obstructive Lung Disease)

stage	severity	$FEV_1\%$
stage 1	mild (GOLD 1)	≥ 80
stage 2	moderate (GOLD 2)	$50 - 79$
stage 3	severe (GOLD 3)	$30 - 49$
stage 4	very severe (GOLD 4)	≤ 30

However even with 30% damaged lung due to emphysema, the spirometry could give misleading normal diagnosis (Lynch et al., 2000). Also subtype detection and damaged lung tissue qauntification or their spatial distributions can not be analysed using spirometry.

8.2.2 Pulmonary CT Based Analysis

Medical imaging is an invaluable tool in medicine for pathology diagnosis. Magnetic Resonance Imaging (MRI), CT, digital mammography, and other medical imaging approaches provide an effective means for non-invasive mapping of the anatomy of a subject. Medical imaging has numerous advantages such as early detection of pathologies, distribution analysis of the diseases, subtype and different pathology detection, quantification of tissue volumes, study of anatomical structure, treatment planning and carrying out computer integrated surgery.

CT imaging, also referred to as Computed Axial Tomography (CAT) scan, is a main imaging tool for diffuse lung diseases. CT imaging has the unique ability to offer clear images of different types of tissue. When the CT slice thickness is also involved, the 3D unit area is known as a voxel. Voxels in a CT scan are displayed in terms of relative radiodensity. The voxel is represented according to the mean attenuation occurred in proportion to the density of the tissue in a 3D unit area and is measured using the Hounsfield units (HU). Water has an attenuation of 0 HU, while air is -1000 HU (Wikipedia, 2014).

The advantages of CT in pulmonary imaging are numerous. Firstly it is fast, which is important for patients who have trouble holding their breath. It is painless, non-invasive, accurate and has been shown to be a cost-effective imaging tool. Also CT scans provide clear images of lung structure which mainly contains air and give high contrast images compared to MRI which is more suitable for more liquidized or solid state tissues and body structures.Furthermore, it can be performed with a patient with implanted medical devices of any kind inside the body, unlike MRI. High Resolution CT (HRCT) is an improved technique of CT, which comparatively produces a large amount of anatomical information which is very much similar to the gross pathological

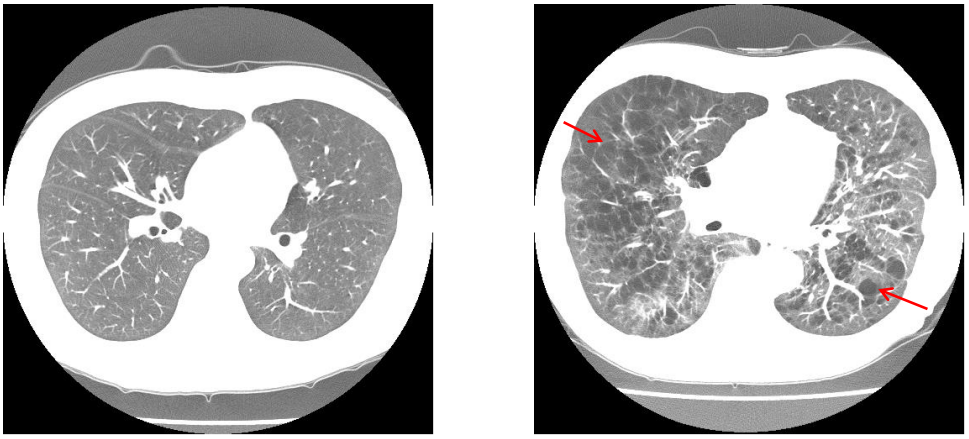


Figure 8.4: HRCT data (a) a HRCT slice from a normal lung, (b) a HRCT slice from a emphysematous lung. Data from (Hara et al., 2004; Sørensen et al., 2013).

specimens or lung slices (Webb et al., 2008). Compared to CT, HRCT produces thinner slices and therefore offers a higher axial resolution.

These advantages of CT imaging make it ideal for evaluation of parenchymal diseases (Sluimer et al., 2003). CT image analysis is a sensitive method for diagnosing emphysema, assessing its severity, and determining its sub types. In pulmonary emphysema, the major advantage of CT analysis over the spirometry approach is that in addition to providing data concerning overall disease severity, it also identifies the specific locations in the lung where the alveoli walls have been destroyed. Furthermore, CT analysis can be also used in sub types analysis (Madani et al., 2001).

In CT analysis, emphysema is characterized by presence of areas of abnormally low attenuation which can be easily contrasted from surrounding normal tissue. Figure 8.4 shows two HRCT slices illustrating the difference between normal and emphysema lung. In severe emphysema low attenuated areas which represent the lung destruction can become confluent and a reduction of blood vessels is clearly visible.

There are three common sub types of emphysema based on anatomical distribution of area of lung destruction. They are Centrilobular, Panlobular and Paraceptal emphysema (Sørensen et al., 2010).

Centrilobular emphysema typically begins near the centre of the secondary pulmonary lobule in the region of the proximal respiratory bronchiole. This is usually seen as small round black (low attenuation areas) evenly distributed holes with ill defined borders that may appear in the central portion of the secondary pulmonary lobule (Fig 8.5a). The resulting area of destruction surrounded by normal tissue allows identification of key structural lesions. This is the most common subtype of emphysema.

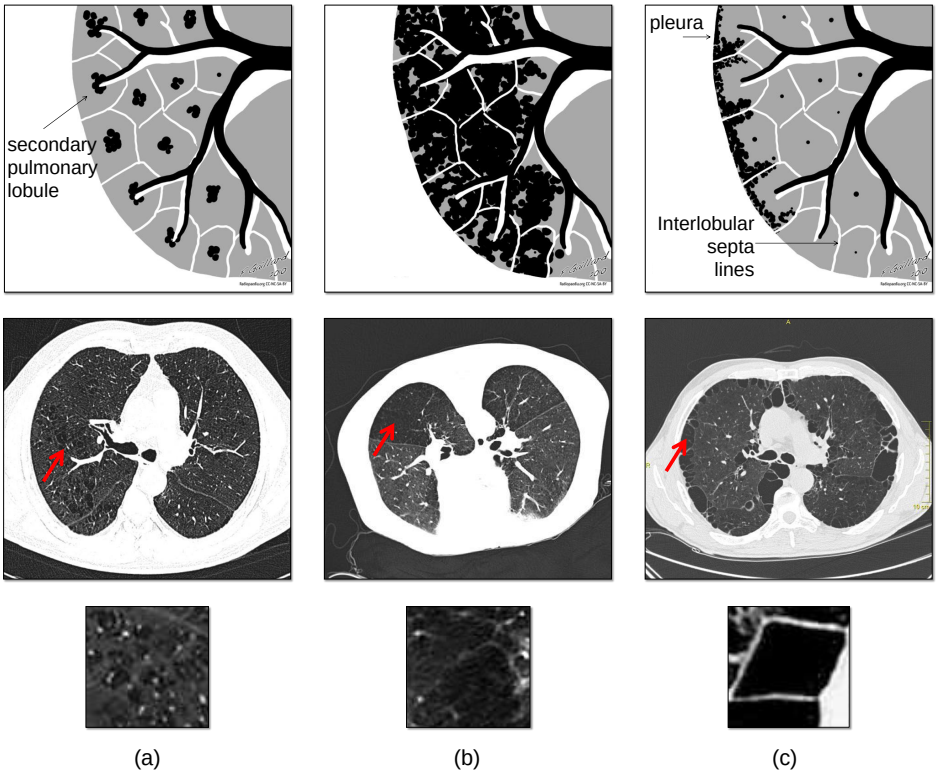


Figure 8.5: Subtypes of emphysema (a) Centrilobular emphysema, (b) Panlobular emphysema, (c) Paraceptal emphysema. Images from Blackmore et al. (2014).

Panlobular emphysema also known as panacinar emphysema typically involves the entire lung and pathologically produces uniform enlargement of all air spaces. The entire lung appears darker with reduction of blood vessels markings. Lung volumes are increased and distinct spaces of low attenuation may not be seen (Fig 8.5b). This subtype associates with the AAT deficiency.

Paraceptal emphysema is located adjacent to the pleura (outermost tissue layer of lung) and interlobular septa lines with a peripheral distribution within the secondary pulmonary lobule (Fig 8.5c). It is visible as multiple low-attenuation areas in a single layer along the pleura often surrounded by interlobular septa that is visible as thin white walls (Fig 8.5c).

Due to the increased amount of information presented in HRCT, manual CT analysis is a very expensive and time consuming task even for experienced radiologists. Therefore, Computer Aided Diagnosis (CAD) of CT scans is widely preferred (Litmanovich et al., 2009). CAD in emphysema analysis mainly depend on the texture features and the intensity information present on the CT scans.

8.3 Current Methods of Computer Aided Emphysema Analysis

CAD approaches based on lung CT scans for emphysema quantification can be categorized into intensity (CT density) based methods and texture based techniques. Under intensity based techniques, many studies have been carried out to find an absolute threshold value below which emphysema is considered to be present. Muller et al. (1988) introduced a method known as 'emphysema index' or 'density mask' which measures the relative amount of lung parenchymal pixels that have attenuation values below -910 HU. Further studies have obtained the optimal threshold value for emphysema quantification as -950HU (Litmanovich et al., 2009). An adaptive thresholding method that also incorporates the information of pixel locations on the lung has been introduced by Hara et al. (2004). Another approach has been proposed based on mean lung density and voxel dimensions (Coxson et al., 1999).

Intensity based CAD techniques usually suffer from problems caused by monotonic intensity changes. This may occur due to several reasons, for example, the influence of contrast materials and degree of inspiration of the lung while scanning (Muller et al., 1988). Secondly there is a degree of corruption in the CT image due to noise, arising from strict limitations on radiation power that can be applied on a patient (Sprawls, 1995). Techniques solely based on intensity are highly sensitive to the noise. Häme et al. (2013) proposed a hidden Markov measure field model to obtain more promising emphysema index measures than standard densitometric approaches showing robustness to noise resulting from reconstruction kernels. However, intensity and texture based integrated approaches are the better choice for pulmonary lung disease analysis using CT (Sørensen et al., 2010).

Texture based CAD techniques are a successful methodology to use in assessing the presence and distribution of emphysema and its subtype patterns. The texture features can be categorized as statistical, spectral and structural features (Litmanovich et al., 2009). Uppaluri et al. (1999) used a method known as adaptive multiple feature method (AMFM) which assessed 22 independent statistical features in order to classify different lung tissue patterns. This approach is further improved by extending it from 2D to 3D by Xu et al. (2006), to classify emphysema and early smoking-related pathologies. Mishima et al. (1999) attempt to detect early emphysema on the basis of fractal analysis. Spatial Gray Level Dependence Method (SGLDM), Gray Level Run Length Method (GLRLM) and Gray Level Difference Method (GLDM) have also been used as statistical feature extraction methods for emphysema diagnosis (Vasconcelos et al., 2010). Depeursinge et al. (2010) have used the density histogram and quincunx wavelet frame coefficients with number of pixels belongs to the air component as an additional feature to evaluate emphysema classification performance with different classifiers. In Kim et al. (2009)

statistical texture features as well as shape features are also employed in lung disease classification.

Sluimer et al. (2003) have used Gaussian, Laplacian of Gaussian and first and second order derivatives of Gaussian filters in different scales and have obtained the histogram of responses as spectral texture features. Depeursinge et al. (2007) have used Discrete Wavelet Frame (DWF) to classify lung tissue types including emphysema. However, filter based methods require selection of an optimal filter set and employing reasonable sizes for filter kernels.

Structural features encapsulate information on structures of the texture such as arrangements of texture primitives. A texton based method has been introduced in Gangeh et al. (2010). Sørensen et al. (2010) applied rotational invariant local binary patterns (LBP) for successful emphysema quantification. This method acquires very compelling results and can be considered as the current state of the art texture based emphysema quantification method.

8.4 Results and Discussion

The application of model based statistical features to emphysema quantification is not commonly used because of the relatively high computational cost. Also the model parameters of TGMRF descriptors are spatially constant, therefore, subtle pattern variations which are crucial in tissue discrimination may be smoothed out. Here we employ I-LPH descriptors to extract tissue texture features because it is comparatively simple and fast and ideally suited for isotropic texture representation such as lung tissue. The small model estimation of IGMRF results in spatial variations in the parameter estimates which closely capture structural arrangement patterns in the texture in addition to the spatial interactions. Therefore, the local parameter distribution based methods integrate the essence of both the statistical and structural characteristics of a texture which makes it more suitable for lung tissue classification task.

8.4.1 Dataset

To evaluate the I-LPH descriptor performance on emphysema diagnosis and quantification we use the online emphysema dataset from Sørensen et al. (2013) which has been also used in performance evaluation of LBP and filter based features (Sørensen et al., 2010). The database comprises 115 HRCT slices of size 512×512 and 168 of square patches of size 61×61 obtained from a subset of slices. The HRCT slices belong to a study group of 39 subjects including non smokers, smokers and smokers with COPD.

Each slice is labeled according to the leading pattern of interest and severity by an experienced chest radiologist and a pulmonologist. The leading patterns are Normal Tissue

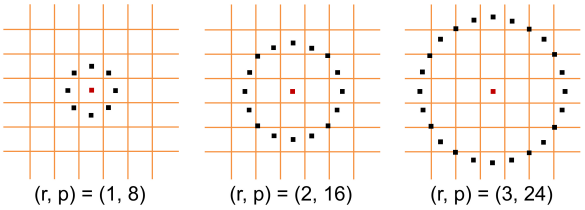


Figure 8.6: Multi-resolution circular neighbourhood system corresponding to $(r, p) = \{(1, 8), (2, 16), (3, 24)\}$.

(NT), Centrilobular Emphysema (CLE), Paraseptal Emphysema (PSE) and Panlobular Emphysema (PLE). The severity of each slice is classified as no emphysema (0), minimal (1), mild (2), moderate (3), severe (4) and very severe (5). The Leading pattern of each patch is also available and there are 59 NT patches, 50 CLE patches and 59 PSE patches. In this study, only NT, CLE and PSE classes are used. However, clinical test results on PFT tests are currently unavailable for the dataset. Therefore, correlations of the results with the diagnosis via visual inspection by experienced radiologists and the emphysema index are considered.

First, the patch dataset is used in a classification framework to identify the discriminative ability of the texture features. Sørensen et al. (2010) reported satisfactory classification performance for this dataset with joint LBP using the parameter setting $(r, p) = \{(1, 8), (2, 16)\}$ and a region of interest (ROI) of size 31×31. The joint LBP features are constructed from 2D histogram of joint intensity-LBP histogram.

8.4.2 Emphysema Classification

Following Sørensen et al. (2010) we also use circular neighbourhoods similar to theirs. The only difference compared to the circular neighbourhoods we used in the previous chapters for I-LPH descriptors is that the number of neighbours p increase with each level. This type of neighbourhood system is suggested by Ojala et al. (2002) for constructing LBP which has been adapted by Sørensen et al. (2010). The name given to these neighbourhood systems is multi-resolution circular neighbourhood systems and is illustrated in Figure 8.6. Bilinear interpolation is used to estimate the neighbour values at off grid positions similar to Ojala et al. (2002) and Sørensen et al. (2010).

According to Sørensen et al. (2010), the present study also employs 31×31 ROIs extracted from each 61×61 patches. For histogram calculation $bins = \{10, 20, 30, 40, 50\}$ are evaluated and $bins = 40$ is selected in the following experiments. The α parameter bin range is set from -1 to $+1$. The σ parameter bin range is set from 0 to 100.

The leave one subject out classification technique discussed in Sørensen et al. (2010) with nearest neighbour classifier is used. In leave one subject out classification, in a trial all the patches belonging to one subject are used as the test set and the rest of the

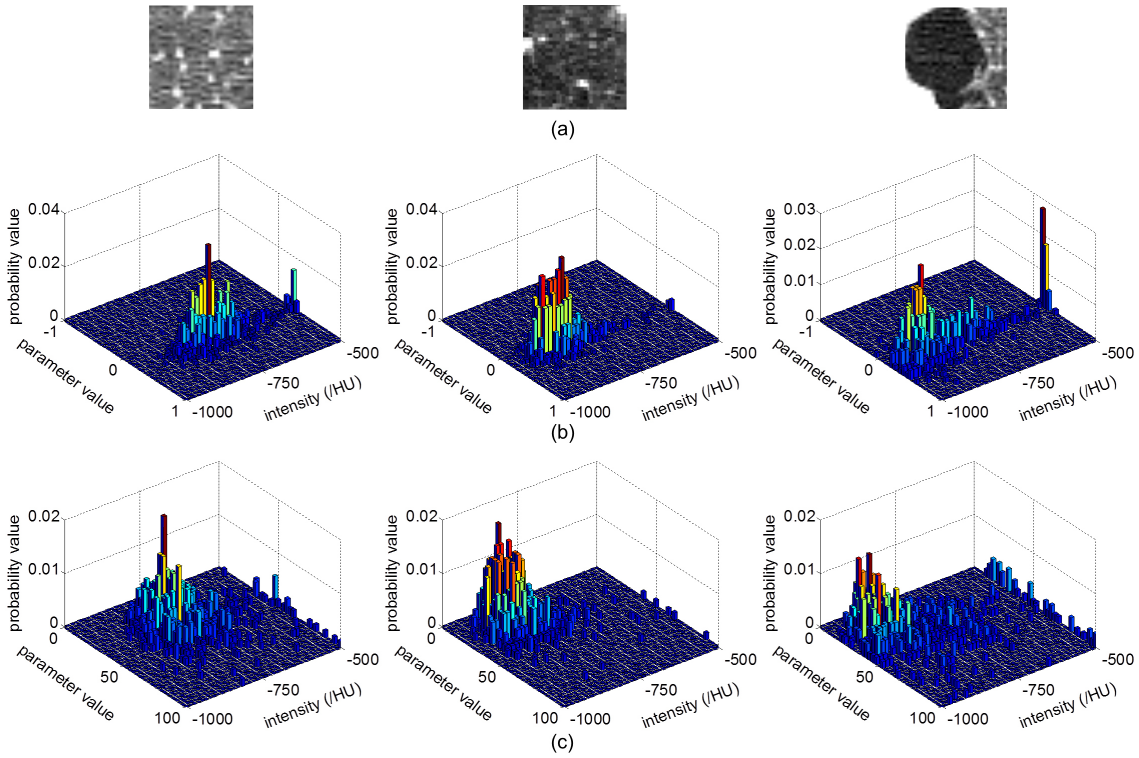


Figure 8.7: 2D intensity I-LPH parameter joint histograms. (a) original patches from the three classes NT, CLE, PSE respectively (31×31), (b) α and intensity joint histograms , (c) σ and intensity joint histograms. $(r, p) = (1, 8)$ and $bins = 40$ in each axis are used.

patches are used as the training set. In each leave one subject out trial, assigned labels (labels assigned by the classification process) are stored. In the end of all trials, the assigned labels are matched against the true labels to calculate the accuracy (Sørensen et al., 2010). The absolute sum of difference between histograms is taken as the distance metric.

Initially, the performance of I-LPH features with and without integrating intensity information is examined. The technique in which the local parameter histograms are constructed without integrating intensity information is referred to as '*noInt*' and the joint intensity-parameter histograms are represented by '*joint*'. Note that '*noInt*' and '*joint*' features are based on I-LPH features and not on LPH features. Figure 8.7 shows the *joint* features of α and σ parameters of IGMRF for some sample ROIs.

Figure 8.8 shows the classification accuracies obtained for the three class problem, NT, CLE and PSE for *noInt* and *joint* features with various neighbourhood systems (r, p). The accuracy gradually decreases with the increasing level of the neighbourhood system suggesting features from lower resolutions, which are close to the centre pixel, are more significant. The case '*subset*' represents the integrated features from multi-resolution levels of the neighbourhood systems $(r, p) = \{(1, 8), (2, 16)\}$.Based only on texture information, an accuracy of 77.4% can be obtained for the case '*subset*' using *noint* features.

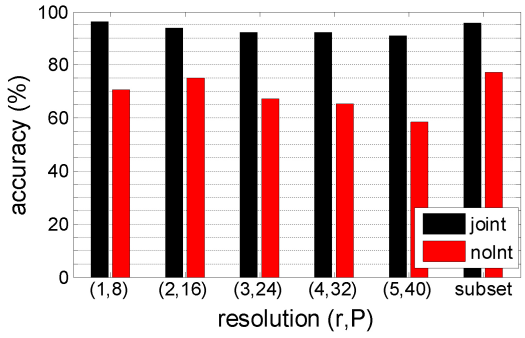


Figure 8.8: Leave one subject out classification accuracy for three class problem, NT,CLE and PSE.

Table 8.2: Comparison with other texture features. Leave-one-subject out classification is used.

descriptor	accuracy (%)
joint(for case subset)	95.8
LBP(for case subset)[1]	95.2
GFB2[1]	94.0
ICR[1]	89.3
INT[1]	87.5

[1] LBP: joint 2D LBP and intensity histogram, GFB2: Gaussian filter response histograms, ICR:Intensity, co-occurrence and run-length, INT: intensity histogram; results reported in Sørensen et al. (2010).

Table 8.3: Confusion matrices for three class classification problem involving the classes NT, CLE and PSE. *LBP* result is obtained from Sørensen et al. (2010).

noInt

	NT	CLE	PSE
NT	37	19	3
CLE	10	40	0
PSE	4	2	53

joint

	NT	CLE	PSE
NT	57	2	0
CLE	2	48	0
PSE	1	2	56

LBP

	NT	CLE	PSE
NT	55	0	4
CLE	1	49	0
PSE	2	1	56

However, integrating intensity vastly improves the accuracy to 95.8%. This is comparable with the accuracies reported for LBP features which is 79.2% with only texture information and 95.2% with joint intensity-LBP features (Sørensen et al., 2010). Table 8.2 illustrates classification accuracies achieved using different texture features. *joint* features perform slightly better than state-of-the-art features, the LBP.

The confusion matrices achieved for the three class problem by the I-LPH descriptors and LBP method are shown in Table 8.3. The LBP results are taken from the joint intensity-LBP histogram performance reported in Sørensen et al. (2010). The confusion matrix of *noInt* features clearly indicate that NT and CLE classes have higher number of misclassified ROIs as a consequence of disregarding intensity information. The *joint*

Table 8.4: Comparison with other GMRF based texture features. Leave-one-subject out classification is used. Note: the *joint* features here are obtained only from $(r, p) = (1, 8)$ resolution.

descriptor	descripton	accuracy (%)
joint	joint intensity & I-LPH histogram	94.6
LPH	joint intensity & LPH histogram	94.0
TGMRF	concatenated with intensity histogram	62.5

features and joint LBP-intensity features perform comparatively well in discriminating different lung tissues.

Next the performance with other GMRF based texture features is considered for comparison. Table 8.4 shows the classification accuracies for *joint* features with $(r, p) = (1, 8)$ and $bins = 20$, LPH features with $n = 3$ square neighbourhood size and $bins = 20$ and TGMRF features with $n = 3$, the second order neighbourhood system. Here, the LPH features are integrated with intensity similar to the *joint* features. The TGMRF features are also used in combination with the intensity histogram.

It can be clearly seen that local parameter distributions are more discriminative compared to TGMRF descriptors. The *joint* and LPH features have a dimensionality of R^{2*bins^2} and R^{5*bins^2} respectively. However, the accuracies for *joint* features and LPH features are almost similar implying that for this specific problem IGMRF is sufficient and additional directional information in LPH features have not been of much use.

Therefore, based on classification performances, *joint* feature are a preferable efficient choice for emphysema and its subtype diagnosis. Results could be further improved by feature selection.

8.4.3 Emphysema Quantification

In this section, we perform lung parenchyma pixel classification of 115 CT slices for emphysema quantification from a study group comprising 39 subjects. The features extracted from 31×31 labelled ROI dataset used in section 8.4.2 are employed as the training models. The training models obtained from a subset of PSE ROIs are used. This subset represents moderate to severe PSE and clearly represent reasonably large regions of PSE tissue pattern with minimum confusion with NT or CLE class tissue. These PSE ROIs contain approximately 15% or less near boundary non parenchymal pixels. This setting of ROIs of PSE class can employ the prior information that PSE has high probability of occurrence near the boundary of the lung parenchyma (Sørensen et al., 2010). During the slice pixel classification all the training patches belonging to the subject of the corresponding slice are left out from the training set.

The $(r, p) = (1, 8)$ setting is used with $bins = 20$ to construct *joint* parameter histogram features. The 6 slices out of 115 slices which belong to class PLE are not considered because the patch data is unavailable. The remaining 109 slices with leading patterns NT, CLE and PSE are employed for tissue quantification. Note that the leading pattern of each slice is available prior to the quantification based on the visual inspection of experienced radiologists.

First of all, the local parameter estimates of each lung parenchyma pixel are estimated according the small model estimation. Then a window of size 31×31 is column-wise scanned on the parameter images and the normalized parameter histograms for each pixel is constructed. The histograms are vectorized and concatenated to formulate the I-LPH *joint* features at each pixel.

The pixels outside the lung parenchyma are directly labelled as the background class by thresholding. The thresholding is based on the knowledge that the CT density values of lung parenchyma pixels are usually between $-1000HU$ to $-500HU$ (Coxson et al., 1999).

Hard and soft classification are performed here. The hard classification assigns each pixel a class label depending on the nearest training model to its feature vector. The soft classification finds the probability a pixel belongs to a certain class. We define the following expression to calculate the class probabilities.

$$p(\omega_c/y_s) = \frac{\exp\{-D(h_s, M_{\omega_c})^2\}}{\sum_{c=1}^{C} \exp\{-D(h_s, M_{\omega_c})^2\}} \tag{8.1}$$

where ω_c represents the class and C is the number of classes. h_s is the feature vector of the considered parenchyma pixel and M_{ω_c} is the nearest feature vector in the training set to h_s from class ω_c. $D(.)$ is the sum of absolute difference distance metric.

Tissue quantification is then carried out on each slice. Two measures are obtained for each slice by fusing the results of all the lung parenchyma pixels on it. The relative hard classification accuracy RCA_{ω_c} (Sørensen et al., 2010) gives the percentage of lung tissue belonging to the class ω_c. Based on soft classification probabilities we define the soft relative probability SRP_{ω_c} for a class ω_c according to,

$$SRP_{\omega_c} = \frac{1}{|\Omega_L|} \sum_{s \in \Omega_L} p(\omega_c/y_s) \tag{8.2}$$

where Ω_L is the lung parenchyma area and $|\Omega_L|$ is the number of pixels in the lung parenchyma.

RCA_{ω_c} is a measure about the spatial extent of the tissue from class ω_c. The SRP_{ω_c} measure introduced here is sensitive to the severe tissue damages localized in a smaller area which can not be quantified properly as a higher degree of tissue damage by RCA_{ω_c}.

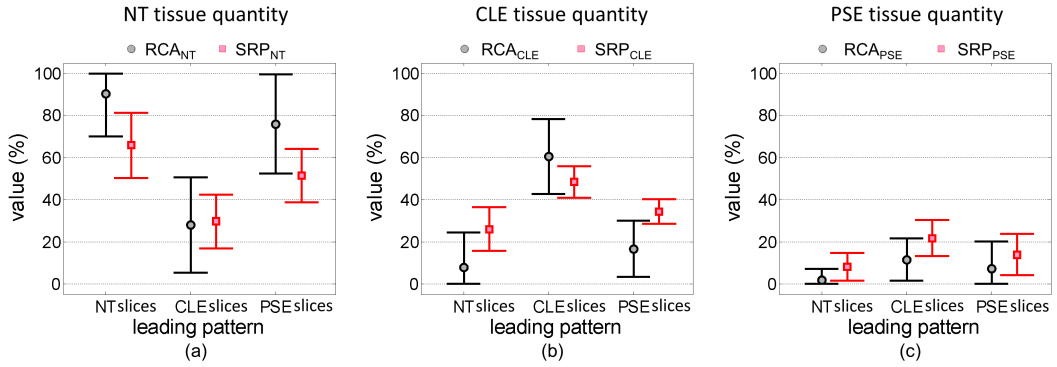


Figure 8.9: Tissue quantification with *joint* features. RCA_{ω_c} and SRP_{ω_c} for each category of slices (NT, CLE and PSE) grouped according to leading patterns labelled by visual inspection judgements. (a) NT tissue quantification, (b) CLE tissue quantification, (c) PSE tissue quantification. *joint* features (2D intensity I-LPH joint histogram) obtained from $(r, p) = (1, 8)$, $bins = 20$ setting are used. Note that the error bars are truncated near 0% and 100%.

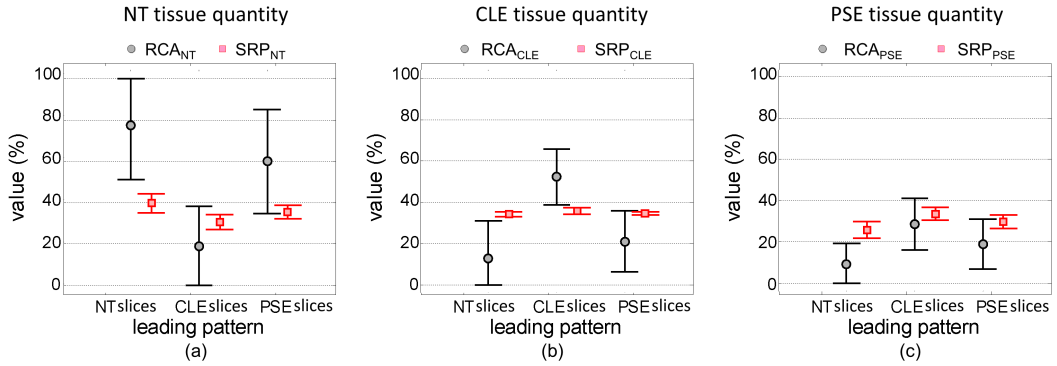


Figure 8.10: Tissue quantification with joint LBP features. RCA_{ω_c} and SRP_{ω_c} for each category of slices (NT, CLE and PSE) grouped according to leading patterns labelled by visual inspection judgements. (a) NT tissue quantification, (b) CLE tissue quantification, (c) PSE tissue quantification. The joint LBP features (2D intensity-LBP joint histogram) obtained from $(r, p) = (1, 8)$, $bin1 = 20, bin2 = 10$ setting are used. Note that the error bars are truncated near 0% and 100%.

The CT slices are categorized into groups according to the leading pattern labellings given by the radiologists. Then the averages and standard deviations of RCA_{ω_c} and SRP_{ω_c} measurements of each category is calculated separately. Results are shown in Figure 8.9 achieved by *joint* features. The $RCA_{\omega c}$ and SRP_{ω_c} counts in Figure 8.9 are consistent with the leading pattern labellings of the CT scans. For example, slices having leading pattern labelled as NT have a higher RCA_{NT} and SRP_{NT} counts and lower RCA_{CLE}, RCA_{PSE} and SRP_{CLE}, SRP_{PSE} counts. i.e. NT labelled lung slices have

more NT tissue quantity (Figure 8.9a error bars under NTslices) and a lower CLE and PSE tissue quantities (Figure 8.9b,c error bars under NTslices). Similarly, CLE labelled lung slices have a higher CLE tissue count and a lower NT and PSE tissue count (Figure 8.9a,b,c error bars under CLEslices). Most of the slices labelled as PSE have minimal or mild PSE except for one severe PSE slice in the considered slice dataset. Therefore RCA_{NT} count is comparatively high indicating less tissue damage extent from PSE in PSE labelled slices (Figure 8.9a error bar under PSEslices). However, it is relatively less than the RCA_{NT} for NT labelled slices.

Furthermore, Figure 8.9c indicates that slices belonging to CLE group also have a tendency to contain higher value of PSE tissue. In the real world situations, CLE patients have a higher chance of co-existence of CLE and PSE. The slices belonging to PSE class have a relatively higher mis-classifications with CLE tissue giving higher RCA_{CLE} and SRP_{CLE} compared to RCA_{PSE} and SRP_{PSE} (Fig 8.9b and c).

The RCA_{ω_c} measures have comparatively larger standard deviations (Figure 8.9) compared to SRP_{ω_c}. This explains the relatively low sensitivity of RCA_{ω_c} for pathology quantification where it only takes into account the extent of the disease but not the strength or the localized severity of the disease.

Figure 8.10 illustrate the NT, CLE and PSE tissue quantifications for labelled slices achieved using joint intensity-LBP histogram features. In comparison to the tissue quantifications achieved by *joint* features illustrated in Figure 8.9, RCA tissue counts from LBP features also show a similar variation among the labelled slices, however with a slightly larger standard deviation. i.e. NT labelled slices have higher NT tissue quantity, CLE labelled slices have higher CLE tissue quantity and PSE labelled slices have higher PSE tissue quantities (see Figure 8.10). Nevertheless the SRP counts obtained from soft classification have a very low inter-class variation (Figure 8.10). This may caused because the dissimilarity between feature vectors from LBP for the classes NTslices, CLEslices and PSEslices is low and it is lower than that of *joint* features (see Figure 8.9 and Figure 8.10). Therefore, more effective and discriminative soft classification results, i.e. SRP measures can be obtained using *joint* features compared to the LBP features.

Figure 8.11 demonstrates some example lung parenchyma images labelled using the proposed approach with *joint* descriptors. The pathology distribution is effectively expressed. Texture is a regional property, therefore, hard classification achieves relatively larger areas of emphysematous tissue than from an intensity based thresholding technique like emphysema index (EI) (Muller et al., 1988). However, the correlation between total emphysema tissue count ($RCA_{CLE}+RCA_{PSE}$) from proposed method and EI with $-910HU$ threshold is large, giving a correlation coefficient of 0.84. Note that EI can not be used in subtype analysis. Moreover, pixels belonging to PSE tissue are often classified near the boundary of the lung parenchyma agreeing to the PSE location dependency

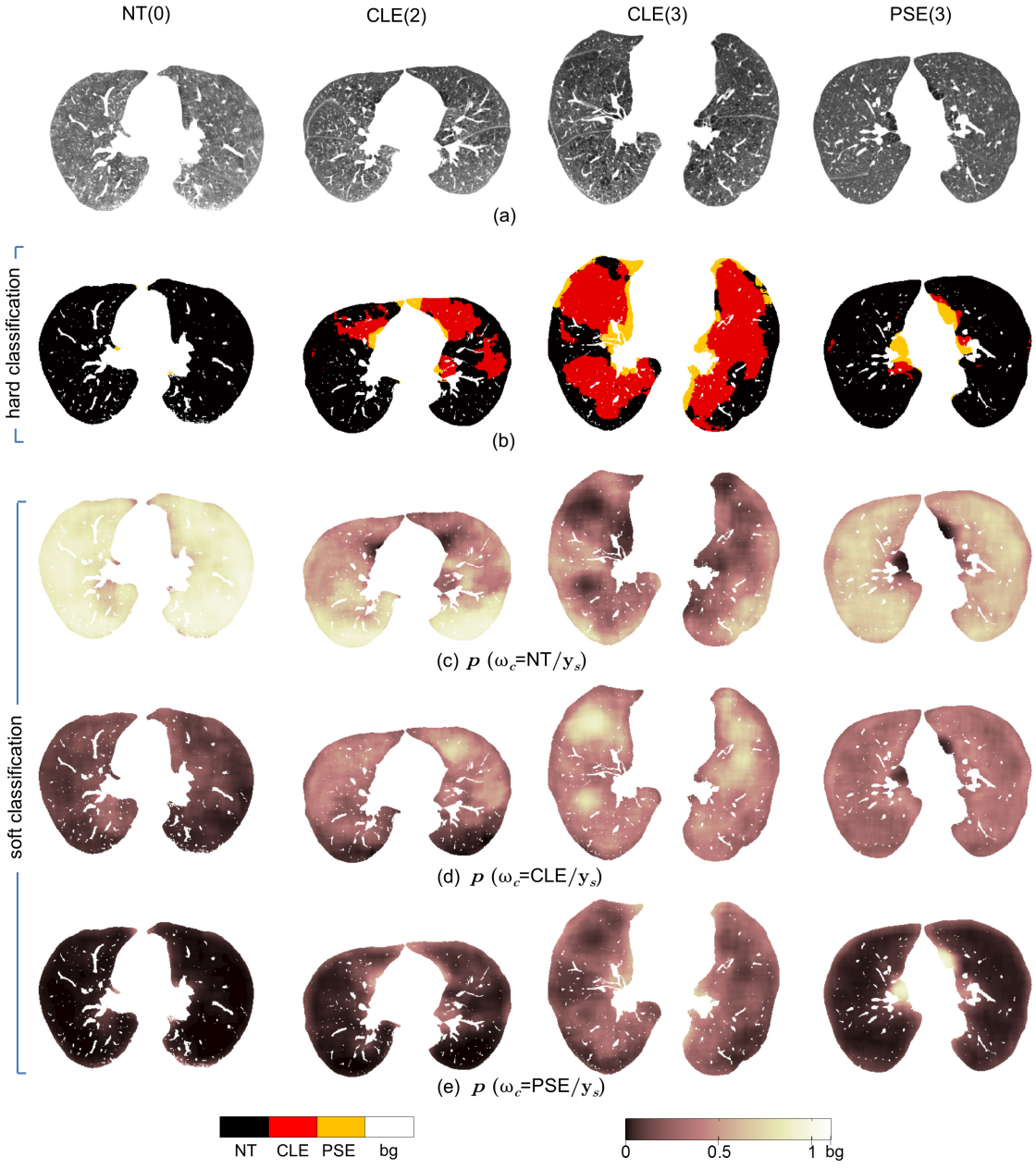


Figure 8.11: Lung parenchyma pixel classification with *joint*. Some example slices representing NT, CLE and PSE slice groups. Severity is indicated in parenthesis. (a) original slices, (b) hard classification results, (c) $p(\omega_c = NT/y_s)$ (d) $p(\omega_c = CLE/y_s)$ and (e) $p(\omega_c = PSE/y_s)$ from soft classification. bg - background (pixels outside the lung parenchyma). *joint* features obtained from $(r, p) = (1, 8)$, $bins = 20$ setting are used.

near boundary (Figure 8.11b). The likelihood that the CLE subjects also have PSE is also clearly reflected in these results.

In comparison to Figure 8.11 achieved using *joint* features, Figure 8.12 illustrates the same lung parenchyma images labelled using LBP features. The hard classification

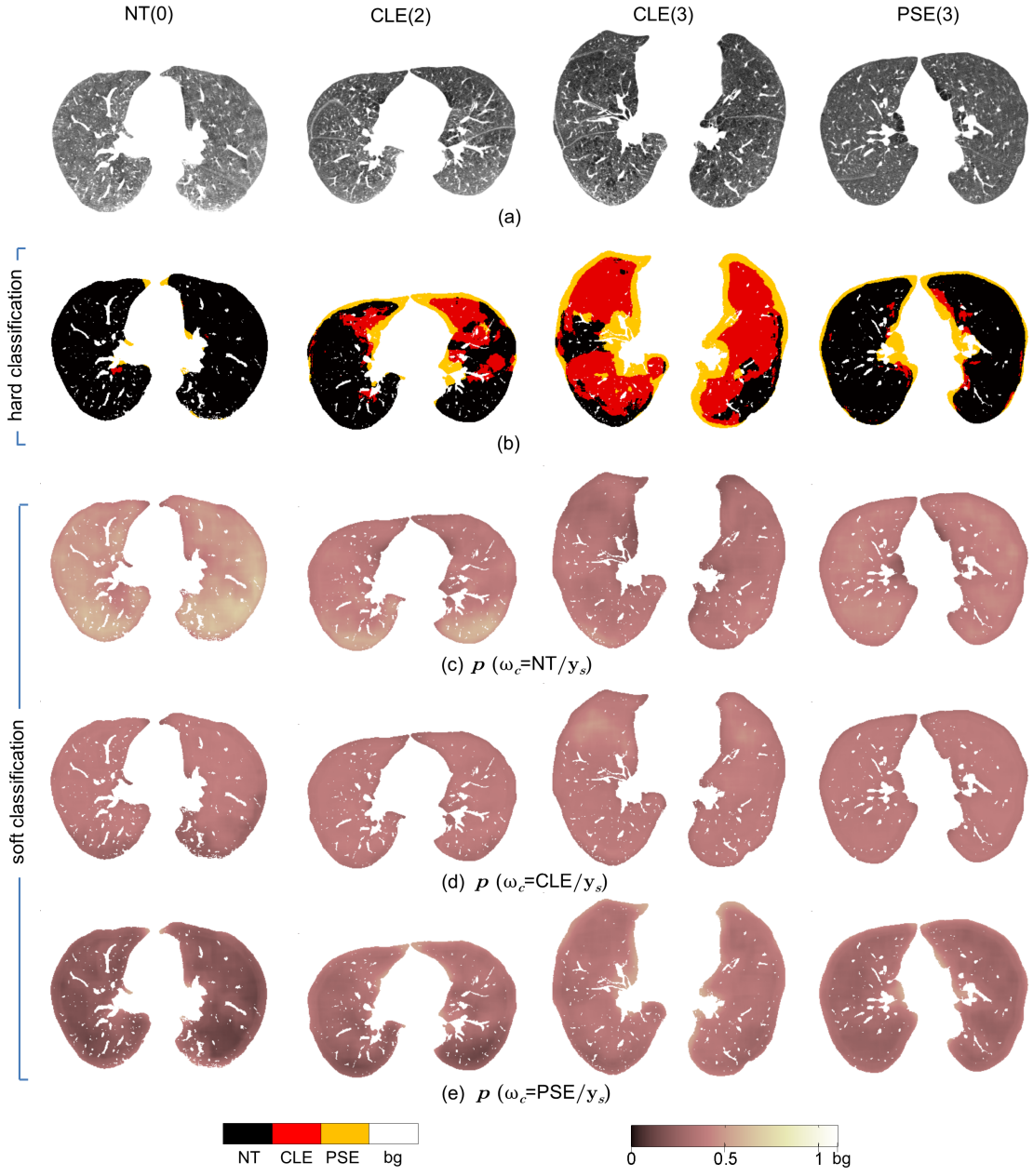


Figure 8.12: Lung parenchyma pixel classification with LBP. Some example slices representing NT, CLE and PSE slice groups. Severity is indicated in parenthesis. (a) original slices, (b) hard classification results, (c) $p(\omega_c = NT/y_s)$ (d) $p(\omega_c = CLE/y_s)$ and (e) $p(\omega_c = PSE/y_s)$ from soft classification. bg - background (pixels outside the lung parenchyma). LBP features obtained from 2D intensity-LBP joint histogram with $(r, p) = (1, 8)$, $bin1 = 20$,$bin2 = 10$ setting are used.

results in Figure 8.12 closely correlate with the results in Figure 8.11, however there is a higher tendency to label the boundary pixels as belonging to PSE tissue than in Figure 8.11. Both features *joint* and LBP perform well in hard classification with good agreement. However, the soft classification results from LBP significantly differ from

that of *joint* features (see Figure 8.11 and Figure 8.12). *joint* features give promising soft classification results relating well with the hard classification results. However, the lower inter-class distances of LBP features between different classes have led to smaller differences between class probabilities of a pixel. Thus, although there are variations of class probabilities according to tissue changes, they are not prominent as the results from *joint* features. Hence *joint* features are more suitable for soft classification and they have a higher discriminative power between three tissue types NT, CLE and PSE compared to LBP features.

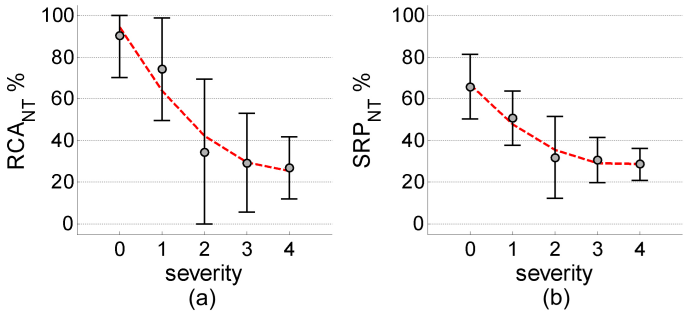


Figure 8.13: Tissue quantification with increasing severity. *joint* features obtained from $(r, p) = (1, 8)$, $bins = 20$ setting are used.

Finally, the severity of the slices and their corresponding tissue percentages are evaluated. The NT tissue counts RCA_{NT} and SRP_{NT} are shown in Figure 8.13. Severity level 5 only has 2 CT slices and is not considered here. The RCA_{NT} and SRP_{NT} measures in Figure 8.13 gradually reduce with increasing severity. i.e. increase in the severity reduces the NT tissue quantity in the CT slices. This implies that more and more normal lung tissues are damaged with the increasing severity of emphysema. Therefore, the results in Figure 8.13 have a good correlation with the severity labellings assigned by visual inspection of the slices.

8.5 Conclusions

In this chapter we have applied I-LPH descriptors to solve a real-world medical problem which is to discriminate and quantify the lung tissue on HRCT data in order to diagnose Emphysema. Emphysema is a common lung disease which fatally disturbs the respiratory process affecting a large proportion of people in the world. I-LPH features are chosen here because they are computationally more efficient and are suitable for classification of emphysema and its subtypes. Results show that these texture features can perform well in discriminating different lung tissues, giving comparative results with the current state of the art texture based emphysema quantification. Furthermore, supervised lung parenchyma tissue segmentation is also carried out. The results illustrate convincing pathology distributions and successful quantification of lung tissues, well correlating with the class and severity labellings by visual inspection. SRP_{ω_c} measure obtained

from soft classification has a higher sensitivity to emphysema pathology quantification than RCA_{ω_c}. Evaluations on other emphysema datasets and correlation to clinical test data need to be considered in future. Also feature selection can be greatly helpful to reduce the higher dimensionality of the features. Nevertheless proposed features are effective for lung tissue classification.

Chapter 9

Bayesian Framework for Integrating Prior Knowledge

9.1 Introduction

In this chapter we formulate a Bayesian framework for LPH construction which has the ability to integrate prior knowledge into LPH features resulting further improved descriptors to be used in future applications. The Bayesian formulation allows prior knowledge about the parameters to be integrated to the local parameter estimation process. In certain applications a priori knowledge about the local texture characteristics may exists which could improve the feature extraction process. For example, in medical image processing, a priori knowledge about more likely locations of existence of a certain disease may known prior to the feature extraction based on human anatomy. Thus, a relationship between the location and the local features could be established using a prior model. Also when handling images contaminated by noise, the smoothing priors could improve the results.

The original formulation of LPH descriptors does not support the integration of such prior knowledge. However, a model which considers prior knowledge in texture feature extraction would be of great use. For the sake of completeness of this research we will introduce a framework capable of integrating prior knowledge into the feature extraction. The prime focus is introducing the concept of the basic framework and empirical study with further developments are left over to the future research work.

Different kinds of prior models can be selected according to the application however, the current study simply employs smoothing prior for the purpose of simplicity in introducing the proposed framework. Any other prior model, if available, could easily replace the smoothing prior in this framework and carry out application specific tasks.

The Bayesian framework discussed here is inspired by the inhomogeneous Bayesian model discussed in Aykroyd (1998) for image reconstruction. However, our study is different from theirs because our objective is constructing an inhomogeneous Bayesian model for texture feature extraction instead of image reconstruction. Therefore, the novelty of this work lies in the likelihood models and prior models introduced here which are capable of characterizing the texture. Thus, the posterior probability model is different from the model in Aykroyd (1998) and is suitable for texture modelling. After the hierarchical formulation of the posterior probability distribution, which is named as the inhomogeneous texture model, the Metropolis-Hastings algorithm is used to estimate the local parameters. Metropolis-Hastings algorithm is used here instead of the local linear regression because it allows finding better parameter estimates with complex posterior distributions that may occur due to the selection of different prior models. The Metropolis-Hastings algorithm is also particularly useful here, because calculating the necessary normalization factor in the posterior model is not required. Finally, distributions of local parameter estimates are constructed using normalized histograms.

9.2 Bayesian Texture Model

In this section we introduce the homogeneous texture model which is subsequently extended to formulate inhomogeneous model for texture. The main difference between the two models is that the homogeneous texture model is defined by spatially constant parameters and the inhomogeneous model is described by spatially varying parameters. The homogeneous texture model is actually the GMRF model discussed in Chapter 3. However, slight change in the notation has been made for consistency and ease of follow up explanations.

9.2.1 Homogeneous Texture Model

Let a stationary random field of a texture on an image region Ω be represented by Y. y_i represents the pixel value at a site i and i is the column-wise linear index. The local conditional model of GMRF describes the relationship between a pixel and its neighbours y_j on a neighbourhood $j \in N_i$ using a Gaussian functional form and is given by,

$$p(y_i|y_j, \boldsymbol{\alpha}, \sigma, j \in N_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y_i - \sum_{j\in\tilde{N}_i} \alpha_j \bar{y}_j\right)^2\right\} \tag{9.1}$$

The $\boldsymbol{\alpha} = [\alpha_j | j = 1 \dots R]^T$ are the interaction coefficients which measure the influence by a neighbour intensity value at the neighbour position j (Petrou and Sevilla, 2006; Li, 2009). R is the number of interaction parameters. The neighbour pixels in symmetric positions about the considered pixel are assumed to have identical parameters (Petrou and Sevilla, 2006), therefore $\bar{y}_j$ is the sum of two neighbour values situated in symmetric neighbour positions with respect to the pixel.

Assuming the conditional independence of a pixel value given its neighbours, the joint distribution can be written as,

$$p(Y|\mathbf{x}) = \prod_i \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} \left(y_i - \sum_{j \in \tilde{N}_i} \alpha_j \bar{y}_j \right)^2 \right\} \tag{9.2}$$

where $\mathbf{x} = [\boldsymbol{\alpha}, \sigma]^T$ is the parameter vector of the model. This will be referred to as the homogeneous model of the texture and it also represents the likelihood of having the texture Y given the GMRF parameter vector. The model parameters of the above model do not depend on the location. Therefore one unique set of parameters will characterize the texture. These spatially constant model parameters are unable to capture the spatial variations in parameters. The homogeneous model, therefore, needs to be modified to describe spatial variations of parameters. The solution is to formulate the inhomogeneous model for texture.

9.2.2 Inhomogeneous Model

The inhomogeneous model is characterized by spatially varying model parameters instead of constant parameters. The local parameter estimates obtained from small model estimation, discussed in Chapter 4 also results in spatially varying parameters. The advantage of using spatially varying parameters is that they can preserve the spatial variations in the pixel interactions acting on the texture.

To obtain spatial variations in parameter space, a separate vector of model parameters for each pixel is defined. In this way, every pixel has its own vector of parameters. Let the parameter vector for pixel at site i be $\mathbf{x}_i = [x_i^j | j = 1, \dots, R+1]$. Note that superscript index j where $j = 1, \dots, R$ represents the type of model parameter according to neighbour position and $j = R+1$ represents the index to the variance parameter. The linear index i represents the location of the pixel similar to section 9.2.1. Hence for every parameter type there will be a corresponding parameter image, $X^j, j = 1, \dots, R+1$, in spatial domain.

Figure 9.1 shows an example to clearly understand the parameters of the model. Here the first order neighbourhood system is considered. Therefore three types of model parameters are involved in characterizing the model, namely, horizontal and vertical

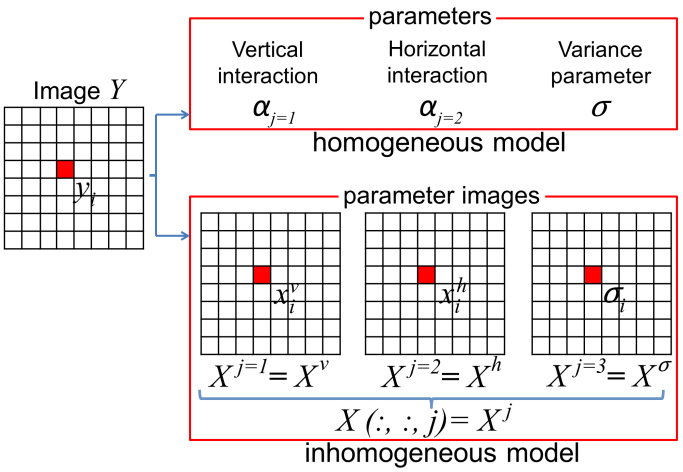


Figure 9.1: Parameters associated with the GMRF models.

interaction parameters and the variance parameter. In the inhomogeneous model, for each type of parameter, for example the vertical interaction parameter, the parameter image is given by X^v which represent the vertical interaction parameter values x_i^v at each pixel $i \in \Omega$ of the spatial domain.

Once the parameters are defined and are assumed to be known, then the likelihood of the texture image Y can be written as,

$$p(Y|X) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{ -\frac{1}{2\sigma_i^2} \left(y_i - \sum_{j \in \tilde{N}_i} x_i^j \bar{y}_j \right)^2 \right\} \tag{9.3}$$

Here the formulation of the inhomogeneous model involves many model parameters compared to spatially constant model formulation. But the estimation process can be easily parallelized using the coding scheme for a much faster estimation process (Petrou and Sevilla, 2006).

The model parameter values on a parameter image will be repetitive according to the pattern repetition. Thus, the unique model parameter values on the parameter image correspond to the parameter values on one texel element. However, the size of the texel is not clearly identifiable in many types of textures, therefore we have considered a region of the texture, Ω as above. Then the distribution of the repetitive model parameters can be used to formulate the texture features.

Next we look at the prior distribution of the model. The prior model on parameters can also be defined as a GMRF on the parameter space. Here we limit our focus to smoothing priors. Alternately, any prior knowledge available on the location dependence of parameters could be associated to the prior model. The prior model for interaction

parameters can be written as,

$$p(X^j|\gamma) = \prod_i \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left\{ -\frac{1}{2\gamma^2} \left(x_i^j - \frac{1}{|N_i|} \sum_{r \in N_i} x_r^j \right)^2 \right\} \tag{9.4}$$

N_i is the neighbours of site i on the parameter image X^j. $|N_i|$ is the number of neighbours. The above prior model is defined for $j = 1, \ldots, R$. i.e. for interaction parameters. γ^2 is the hyper variance parameter of spatially varying model parameters. It is considered that the value of γ is same for any interaction parameter X^j for $j = 1, \ldots, R$. The prior model for the variance parameter is,

$$p(X^{R+1}|\delta) = \prod_i \frac{1}{\sqrt{2\pi\delta^2}} \exp\left\{ -\frac{1}{2\delta^2} \left(\sigma_i - \frac{1}{|N_i|} \sum_{r \in N_i} \sigma_r \right)^2 \right\} \tag{9.5}$$

where δ is a constant which represents the hyper variance parameter of spatially varying σ_i parameter. σ_i is same as the x_i^{R+1} and is used for better readability (Figure 9.1).

9.2.3 Bayesian Formulation

The posterior density for the inhomogeneous texture model can be written as follows.

$$\begin{aligned} p(X, \gamma, \delta|Y) &= \frac{p(X, \gamma, \delta, Y)}{p(Y)} \\ &= \frac{p(Y|X, \gamma, \delta)p(X|\gamma, \delta)p(\gamma, \delta)}{p(Y)} \end{aligned}$$

We assume the conditional independence between various variables to simplify the above expression, including the independence between hyper parameters, γ and δ. The posterior density can be then written as,

$$\begin{aligned} p(X, \gamma, \delta|Y) &= \frac{p(Y|X)p(X|\gamma, \delta)p(\gamma)p(\delta)}{p(Y)} \\ &= \frac{p(Y|X) \prod_{j=1}^{R} p(X^j|\gamma)\, p(X^{R+1}|\delta)p(\gamma)p(\delta)}{p(Y)} \end{aligned} \tag{9.6}$$

Since there is no prior knowledge about γ and δ uniform distributions for $p(\gamma)$ and $p(\delta)$ are assumed. All the other densities are defined as in section 9.2.2.

It is important to mention about the local conditional models of the joint models in equations (9.3), (9.4) and (9.5). A local conditional model tells us about how a pixel depends on its neighbours. Even though we use the global models in MCMC estimation,

when calculating acceptance probabilities, all the terms will cancel out, except terms associated to the local models due to Markovian property given a symmetric proposal distribution. The local conditional models of joint models in equations (9.3), (9.4) and (9.5) intuitively become the expressions without the product symbol. But in this study we use a slightly altered local model for (9.3) as below.

$$p(y_i|X_i) = \prod_{r}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left\{ -\frac{1}{2\sigma_i^2} \left(y_r - \sum_{j \in \tilde{N}_r} x_i^j \bar{y}_j \right)^2 \right\} \tag{9.7}$$

m represents the number of immediate neighbours of site i, for example eight neighbours around i. This expresses that probability of a pixel value not only depends on its neighbourhood but also on near by m pixels and their neighbours. It is a localized version of the likelihood of m samples. Therefore σ_i^2 represents the variance considering m local samples which is a localized variation at site i. Prior to using m local samples their sample mean is set to zero.

9.2.4 Parameter Estimation

Following Aykroyd (1998) we also use the Metropolis-Hastings (MH) algorithm for parameter estimation. Therefore, finding normalizing constant of posterior distribution in (9.6) is no longer needed. Each type of parameters including γ and δ are sequentially estimated in turns. Approach to the estimation of various groups of model parameters is the same.

Let model parameters be represented by Θ where $\Theta = \{X, \gamma, \delta\}$. Let the parameter being considered be θ_i. A proposed new value is selected from the proposal distribution $q(\theta_i^{'}|\theta_i)$. The set of parameters containing the proposed value is given by $\Theta^{'} = \{\theta_1, \ldots, \theta_{i-1}, \theta_i^{'}, \theta_{i+1}, \ldots, \theta_{|\Omega|(R+1)+2}\}$. The proposed value of parameter is accepted and then updated with the acceptance probability,

$$\min\left\{ 1, \frac{p(\Theta^{'}|Y) q(\theta_i^{'}|\theta_i)}{p(\Theta|Y) q(\theta_i|\theta_i^{'})} \right\} \tag{9.8}$$

Otherwise it is rejected and the previous value is retained. Doubly exponential distribution centred on the current value is used as the proposal distribution. The scale parameter of the proposal distribution $q(\theta_i^{'}|\theta_i)$ is chosen by trial and error technique. Since the proposal distribution is symmetric the ratio $q(\theta_i^{'}|\theta_i)/q(\theta_i|\theta_i^{'})$ in (9.8) is cancelled out. The values of X^{R+1}, γ and δ are chosen to be positive all the time.

Many terms of the ratio in (9.8) will cancel out due to Markovian property. This leads to vastly simplified expressions. To avoid numerical overflow, the log value of posterior

ratio is used. The Markov chain is developed with the accepted samples chosen according to the acceptance probability. The convergence of the chain is monitored graphically. When the chain is converged the average of samples laying outside the burn-in period are used as the expected value of the corresponding parameter.

Once the model parameters are estimated in this way, their spatial distributions constructed by normalized histograms can be used to formulate discriminative texture features.

9.3 Results and Discussion

The Bayesian framework for the textures proposed here can be used to extract spatially varying model parameters and their spatial distributions can be used as effective texture features for classification.

In this study, the focus is limited to the first order neighbourhood system of GMRFs. Therefore, three different types of model parameters, namely horizontal interaction parameter, vertical interaction parameter and variance parameter characterize the model.

Spatially varying model parameters are estimated by sampling the proposed posterior probability distribution in (9.6) and then taking the expected values of the samples excluding the burn-in period. MH algorithm is performed on each individual site to sample and estimate the parameters at that location. The coding scheme is used to speed up the process where instead of visiting each site in the image sequentially, a batch of pixels belonging to the same code is updated in parallel (Petrou and Sevilla, 2006).

Each Markov chain is run for 2000 iterations. The first 500 samples are considered as the burn-in period. Rest of the samples are used to calculate the expected value of the parameter.

The scale parameters of doubly exponential proposal distributions are set by trial and error method for each type of model parameters. For Markov chain updates of interaction parameters, $X^j, j = 1, \ldots, R$ the scale parameter is 0.05 and for the variance parameter X^{R+1} it is 0.1. For super parameters γ and δ, 0.05 and 0.1 are used respectively. The number of local samples for the likelihood model, m is restricted to the five nearest samples in the proximity of considered site. This parameter setting is kept constant for all the experiments unless stated otherwise.

In Figure 9.2 examples of estimated parameter images (expected values) achieved using the inhomogeneous Bayesian framework are shown. Figure 9.2a has two texture regions and corresponding horizontal and vertical interaction parameter images illustrate the spatially varying nature of the estimated model parameters. The variance parameter

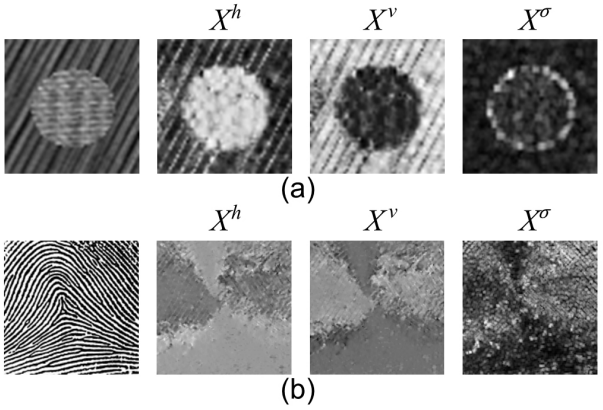


Figure 9.2: Parameter images obtained by inhomogeneous Bayesian framework for parameter estimation. X^h - horizontal interaction parameter, X^v - vertical interaction parameter, X^σ - variance parameter

clearly indicates a higher variance near the boundary between the two texture regions. The pattern inside the circular region in Figure 9.2a has more noteworthy horizontal interactions. The corresponding horizontal interaction parameter image has higher interaction parameter values in respective region. Also the pattern outside the circular region has a directional pattern closer to the vertical axis. Hence the corresponding vertical interaction parameter has higher interaction values outside the circular region.

Figure 9.2b shows a partial finger print. The spatially varying parameters are not made rotational or scale invariant here. Therefore the corresponding spatially varying parameters capture the directional differences in the patterns of the finger print. In general, by looking at Figure 9.2 it can be concluded that spatially varying model parameters carry more information about the texture.

The Markov chains of super parameters γ and δ of Figure 9.2a are shown in Figure 9.3. These chains graphically indicate the convergence roughly after 200 iterations. Thus, the burn-in period and number of iterations mentioned earlier are suitable for the sampling process. However, to perform 1000 iterations to produce a Markov chain of length of 1000 samples for an image of size 256×256 it require 2533 seconds on a a Matlab R2013a environment running on a 2.67 GHz CPU. i.e. approximately 42 minutes. Nevertheless, the speed may be improved by parallel computing which has not been considered here.

The local distributions of these spatially varying parameters can be used to discriminate texture regions. An evaluation of spatial distributions of spatially varying parameters is illustrated in Figure 9.4. Four textures are used and their corresponding spatially varying parameter images are obtained and converted into normalized histograms with 50 bins. The parameter image values are modified to lie between the range 0 to 255 before constructing the histograms. This normalization is done for illustration purposes only. The intensity histograms of the four textures are given in Figure 9.4c.

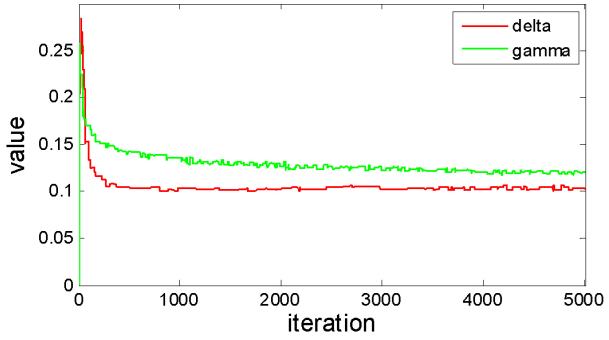


Figure 9.3: Markov Chains of δ and γ parameters for image in figure 9.2a.

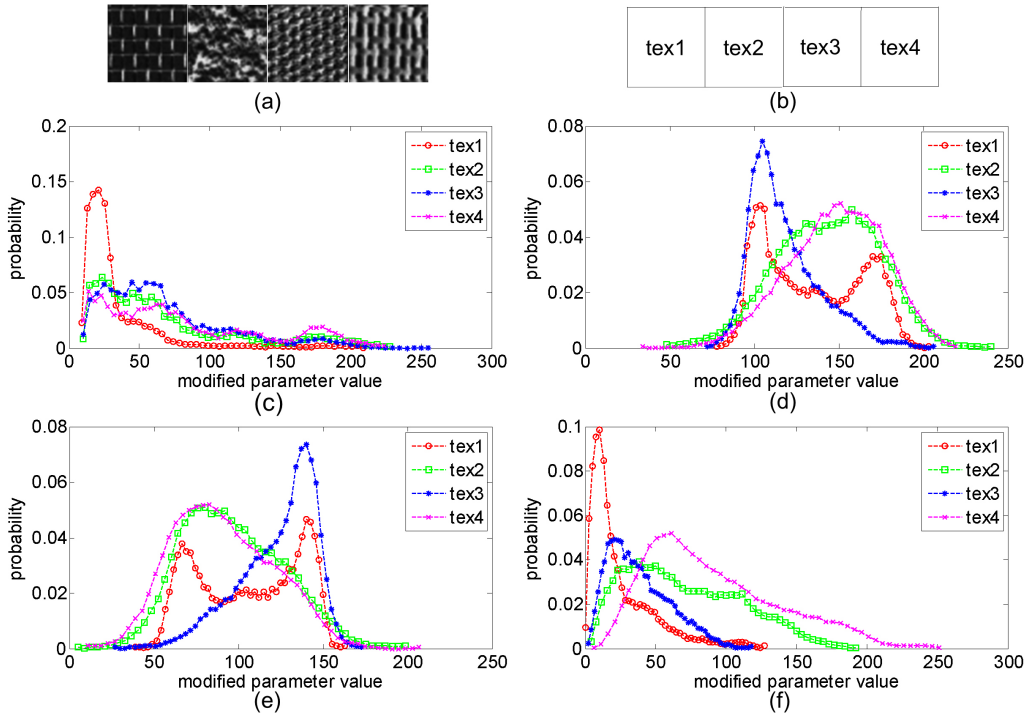


Figure 9.4: Histogram comparison. (a) image comprising four textures, (b) texture labels, and (c) intensity histograms. Histograms of spatially varying parameter images (d) histograms of X^h (e) histograms of X^v (f) histograms of σ of each texture.

According to Figure 9.4c it can be seen that discrimination power of the intensity histogram is quite low for these four textures. But the histograms of parameter images represent substantial differences in there distributions. The interaction parameter histograms in Figure 9.4d and e show a negative correlation between the distributions for a texture. Here we have only two interaction parameters in the model. Therefore, when the horizontal interaction is dominant, vertical interaction of the respective texture is much insignificant. However, these distributions of spatially varying model parameters can be used as a discriminative texture feature in texture analysis.

Table 9.1: Accuracy comparison with other methods. First order neighbourhood system is used. $bins = 50$ is used for *IBMF* and *LPH* methods.

method	method	
	BRODATZ	*OUTEX*
IBMF	88.3 ± 1.57	87.1 ± 1.75
HBMF	37.9 ± 2.19	31.7 ± 2.19
LPH	81.0 ± 1.13	83.1 ± 1.99
TGMRF	40.7 ± 2.02	40.7 ± 2.59
LBP	81.5 ± 2.05	84.8 ± 1.56

9.3.1 General Texture Classification

We perform texture classification using the two datasets used in Chapter 7, namely *BRODATZ*, comprising 32 Brodatz textures (Brodatz, 1966) and *OUTEX*, having 24 OUTEX textures (Ojala et al., 2002). For OUTEX dataset only 20 randomly selected samples per class are used. Prior to feature extraction all the images are pre-processed using histogram equalization. Next, parameter estimation is carried out and the distributions of spatially varying parameters are constructed by normalized histograms.

The classification experiments are performed using equal sizes of training and test datasets randomly partitioned to have equal class proportions. The experiment is repeated 100 times with different training and test sets. Accuracies reported here are the mean accuracy of 100 iterations and its standard deviation. Classification is performed using nearest neighbour classifier with absolute difference distance metric.

The accuracies are given in table 9.1. The proposed inhomogeneous model based feature extraction is referred to as *IBMF* which stands for 'Inhomogeneous Bayesian Model based Features'. The feature extraction based on homogeneous Bayesian model is labelled as *HBMF* which stands for 'homogeneous Bayesian Model based Features'. Four other methods have been used for performance comparison. The LPH descriptor is based on the spatially varying model parameters estimated using small model estimation which is discussed in Chapter 4. However, here only the first-order neighbourhood system is employed in feature extraction. i.e. only horizontal and vertical interaction parameters are presented ($n < 3$). TGMRF is the traditional GMRF feature extraction method with the first-order neighbourhood system (Manjunath and Chellappa, 1991). LBP represents the rotational invariant uniform local binary patterns features (Ojala et al., 2002). Only the LBP histograms from (p=4,r=1) are used to get roughly similar neighbourhood representations as the GMRF setting where the first order neighbourhood system is used. Note that we have only used the first-order GMRFs here. By increasing the neighbourhood size accuracies can be further improved.

It is observed that the *IBMF* has better accuracy compared to LPH features based on local linear regression (table 9.1). A comparison between classification accuracies

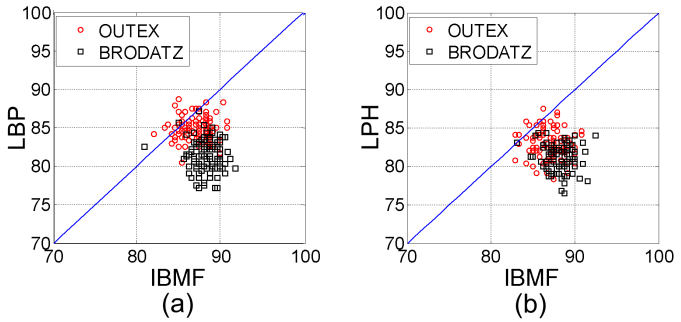


Figure 9.5: Comparison between accuracies obtained from *IBMF* with *LBP* and *LPH* in 100 repetitive classification trials. (a) *IBMF* with *LBP* (b) *IBMF* with *LPH*.

obtained in 100 repetitions of classification experiment using LBP and LPH descriptors with *IBMF* are shown in Figure 9.5. The *IBMF* features perform better in most of the trials than other features (see Figure 9.5).

In the present study, we have employed the local smoothing priors for *IBMF* method. Therefore, by integrating appropriate prior information, *IBMF* features can perform better than the simple least square estimation based *LPH* features.

9.3.2 Robustness to Additive Noise

Here we evaluate the performance of LPH feature with additive noise and how prior knowledge integration can help in this situation. The BRODATZ mosaic dataset used in Chapter 6 for texture segmentation is employed here. Additive Gaussian noise with standard deviation σ_{im} is added to the dataset images and segmentation is carried out similar to Chapter 6 with k-means clustering algorithm. The segmentation error box plots for two situations where $\sigma_{im} = 10$ and $\sigma_{im} = 30$ is shown in Figure 9.6.

According to Figure 9.6 it is observed that when image noise σ_{im} increases the segmentation error s_e increases. However, for *IBMF* where smoothing prior is integrated, the increase in the s_e is less than that for LPH features where no prior knowledge has been assumed. Therefore, by integrating smoothing priors, LPH features can be made more robust to the image noise.

9.4 Conclusions

This chapter considers building a technique to integrate prior knowledge to the LPH descriptors. According to the nature of different vision applications there could be related prior knowledge which would further improve the parameter estimation process.

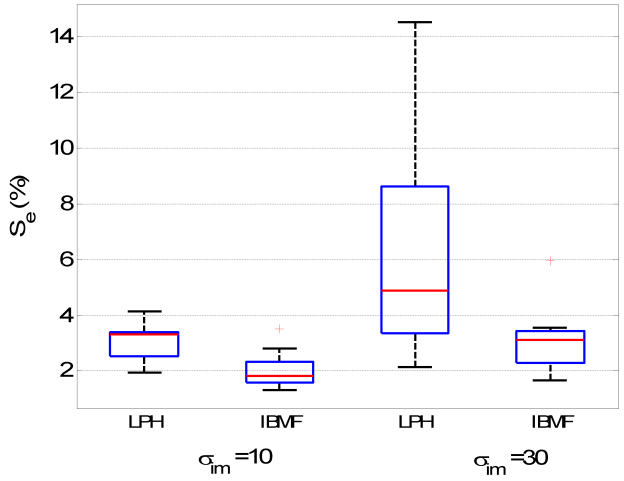


Figure 9.6: Segmentation error s_e for BRODATZ mosaic dataset with additive Gaussian noise using LPH and *IBMF* descriptors.

For example, these priors could be explaining the location dependencies, colour dependencies, smoothness etc. The Bayesian formulation allows integrating prior knowledge about the parameters to the estimation process. The current study uses smoothing priors to locally smooth the spatially varying parameter space. Therefore, this approach can reduce the noise present in the spatially varying parameters while preserving the discriminative ability of the features formulated based on them. Our Bayesian framework for spatially varying parameter estimation is inspired by the inhomogeneous Bayesian model discussed in Aykroyd (1998) for image reconstruction. However, our study is different from theirs because our objective is constructing an inhomogeneous Bayesian model for texture. Therefore, proper likelihood models and prior models are introduced to capture texture characteristics. After formulating the inhomogeneous texture model, the Metropolis-Hastings algorithm is used to estimate the spatially varying parameters. The distributions of spatially varying parameters are then constructed using normalized histograms. The experimental results show that this approach can produce more discriminative texture features in the presence of noise. However, this method needs to be extended and analysed using different priors other than smoothing priors which could show the real value of integrating prior knowledge to the feature extraction process. But currently there are no application specific data available for this experiment therefore, it is left for future research.

Chapter 10

Conclusions and Future Work

10.1 Conclusions

The main focus of this research is developing a novel, improved texture descriptor based on GMRFs and evaluating its performance in texture classification and segmentation. The TGMRF features have been successfully employed in texture classification and segmentation tasks, however some other studies have reported its decreased discriminative performance. Therefore, in this study we have analysed the generalized performance of TGMRF features and examined their weaknesses in order to formulate improved features. In Chapter 3 the main flaws affecting the discriminative quality of TGMRF features are discussed. This evaluation demonstrates that the global parameter estimation process associated with TGMRF parameter estimation results in over-smoothed estimates. These estimates show reduced ability to characterize several important properties of texture, such as structural information and some local statistical properties of the texture. Thus, in Chapter 4 a novel texture descriptor named as LPH descriptor is proposed which can overcome the problems associated with traditional GMRF features. Here, local linear models are fitted in order to characterize the texture, under specific restrictions and modifications. This estimation process is referred to as small model estimation. The small model estimation procedure results in local estimates which are spatially varying.

Local parameter estimates capture the local structure of the texture, or in other words, they model local spatial interactions of a texture. Spatially varying local parameter estimates characterize the partial or complete texture primitives and thus embed the structural information about the texture. The histogram of these local parameter estimates represents the information on the statistical distribution of texture primitives characterized in terms of local parameter estimates. Therefore, LPH descriptor enables

integration of both statistical and structural information about the texture. LPH descriptors overcome the smoothing problem associated with GMRF feature formulation and provide more descriptive texture features.

Methods of achieving rotation invariant texture features are discussed in Chapter 5 and two rotation invariant texture descriptors are introduced, namely RI-LPH and I-LPH descriptors. The performance of these texture features is investigated on texture segmentation and classification and is illustrated in Chapter 6 and Chapter 7 respectively. The texture discriminative ability of LPH, RI-LPH and I-LPH features are significantly better than TGMRF features. Thus, for texture segmentation and classification where more descriptive features are favourable compared to generative features, the distributions of local parameter estimates in the form of LPH, RI-LPH and I-LPH descriptors are highly recommended over TGMRF features. RI-LPH performs relatively better compared to LPH and I-LPH descriptors because it is rotational invariant and includes directional information. I-LPH descriptors on the other hand are more suitable for isotropic texture and provide efficient feature formulation compared to LPH and RI-LPH features. Comparisons with existing state-of-the-art texture features demonstrate that LPH, RI-LPH and I-LPH descriptors have comparable classification results and improved segmentation results. In texture segmentation, integrating colour information can improve the segmentation performance. Furthermore, advanced segmentation methods such as active contours improve the final segmentation results compared to the k-mean clustering algorithm.

The use of I-LPH descriptor on a real world medical application is examined in Chapter 8. Lung tissue classification and quantification for emphysema analysis is performed using lung HRCT data. I-LPH descriptors are selected for this task because they are relatively efficient and adequate for discriminating isotropic textures such as textures of emphysema subtypes. Successful tissue classification and lung parenchyma segmentation results are achieved. The lung tissue classification task involves examination of localized areas and therefore requires localized texture feature extraction processes. Also texture features that provide high discriminative capabilities are required. Thus, distributions of local features such are I-LPH are well suited for this task. Compared to the current state-of-the-art texture based emphysema quantification method, the joint LBP-intensity histogram features, I-LPH descriptors give comparable texture classification performances and improved tissue quantifications and pathology distributions.

Some preliminary work on prior knowledge integration to local parameter estimation of LPH is introduced in Chapter 9. A Bayesian framework is proposed to integrate prior knowledge and the smoothing prior is employed to explain the framework. A smoothing prior can be used to make LPH descriptors more robust to additive noise present in the images. By replacing the smoothing prior with application specific different prior knowledge, the proposed framework can be adapted to integrate different kinds of prior

knowledge, such as location tendencies, colour dependencies etc. These analyses are left for future research.

10.2 Future Work

The current work mainly focus on LPH feature extraction and analysing the performance in classification and segmentation tasks. The following are some of the future directions to be explored that can contribute to further improvements.

- *Texture Synthesis:* Texture synthesis is one of the main fields in texture analysis with many applications. The TGMRF features are generative features which can be directly use in texture synthesis. In TGMRF feature extraction one set of model parameter estimates are acquired and the conditional model with estimated parameters can be employed in texture synthesis via a sampling technique such as Gibbs sampling (Li, 2009). The features proposed here however, are more descriptive and vary from the original form of GMRF features. They have the forms of distributions of local parameter estimates. Therefore, different techniques are required to implement texture synthesis. One plausible suggestion is constructing image pyramids and applying histogram matching. Pyramid based texture synthesis has been introduced by Heeger and Bergen (1995) employing filter bank responses. This technique could be successfully employed with the proposed features for texture synthesis. For example, with LPH features we construct parameter histograms corresponding to different neighbours in different directions and different distances from the centre pixel. These histograms could be used to replace filter response histograms in the pyramid method.

- *Prior Knowledge Integration with Bayesian Framework:* In this research preliminary work on a Bayesian framework to integrate prior knowledge into the LPH feature extraction process was discussed in chapter 9. Therefore, we have introduced an inhomogeneous Bayesian model suitable for characterizing texture which produces spatially varying model parameters. Distributions of spatially varying model parameter estimates are then formulated as the homogeneous texture descriptor. The likelihood and the prior to derive the posterior model for texture modelling are introduced. Here we have used smoothing priors to evaluate the proposed model. However, there may exist many other forms of prior knowledge based on specific applications. For example, in a medical image problem a specific region where a certain disease could occur with a higher probability may be pre-known based on the medical records or human physiology. These location information could be used as the prior knowledge to build a prior probability model. Similarly colour dependencies, i.e. given a certain colour, occurrence of a certain texture could be higher, or any other prior data could exist. Therefore, integrating such prior knowledge into the proposed model should be considered in future. Furthermore, for the parameter estimation we have employed iterative Metropolis-Hastings algorithm

which has become a time consuming process when a large number of model parameters are present in the model. Therefore, methods to improve the efficiency should be considered in future. Nikou et al. (2010) have suggested a faster estimation process in their Bayesian framework with mixture models. Their technique could lead to finding a related but different approach to improve the efficiency in parameter estimation in our proposed Bayesian model.

- *Evaluations on Other datasets:* We have used a publicly available emphysema dataset to evaluate the proposed texture features on emphysema quantification and subtype analysis. However, the performance of these features on other medical datasets is worth consideration. The proposed texture descriptors could be easily applied for analysis on other lung diseases, such as Idiopathic Pulmonary Fibrosis (IPF) which is even more responsive to texture analysis than emphysema. Also these features can be used in other texture based applications such as remote sensing.

- *Integrating Information on Structural Arrangements:* The LPH descriptors are based on constructing histograms on parameter images which are achieved by local parameter estimation. The histogram construction looses the arrangement patterns revealed on parameter images. Therefore, better techniques which can capture the arrangement patterns of local parameter estimates could further improve the performance. However, even without this information, the LPH descriptors still perform well. Techniques to improve the efficiency of feature formulation would also be of great use in order to encourage many studies to select these features in their work.

- *3D Texture Descriptors:* Proposed LPH, RI-LPH and I-LPH features could be extended to 3D texture descriptors to employ in 3D texture analysis. 3D texture descriptors have been perform better than 2D texture descriptors in 3D texture analysis, for example with 3D medical data (Xu et al., 2006).

Appendix A

Integrated Active Contours

Integrated Active Contours is a segmentation technique introduced by Sagiv et al. (2006) for image segmentation including textured images. It is an integrated framework which combines the boundary and the region information in active contour modality. We modify the segmentation process of Sagiv et al. (2006) at the curve evolution step, moving to a technique that increase the performance of curve evolution which is suggested by Mahmoodi (2009). Here, the shape characteristic function is used in evolution process instead of the sign distance function. Using active contours for GMRF based texture segmentation, instead of an optimization via the relaxation methods, provides faster convergence (Mahmoodi, 2009). But with active contours the risk of segmentation being converged on to local minimums increases due to the gradient based optimization (Mahmoodi and Gunn, 2011).

Let C be the evolving contour, and $\mathbf{F}_{in}^k$ and $\mathbf{F}_{out}^k$ be the average PL histogram of inside and outside the contour C. Then the energy functional to minimize is given by,

$$\begin{aligned} E(\mathbf{F}_{in}^k, \mathbf{F}_{out}^k, C) &= \mu \int_{on(C)} h(x,y)\,\mathrm{d}x\,\mathrm{d}y + \lambda_{in} \sum_{k=1}^{(n^2+1)/2} \int_{inside(C)} D(\mathbf{F}_s^k, \mathbf{F}_{in}^k)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad + \lambda_{out} \sum_{k=1}^{(n^2+1)/2} \int_{outside(C)} D(\mathbf{F}_s^k, \mathbf{F}_{out}^k)\,\mathrm{d}x\,\mathrm{d}y \end{aligned} \tag{A.1}$$

where $D(.,.)$ denotes the L1 norm distance between PL histogram features at a pixel and the corresponding average PL histogram. μ, λ_{in} and λ_{out} are constants which determine the contribution of each term to the total energy. The functional in equation (A.1) assumes the texture image consists of two textures. It can be generalized for images with more than two textures by adapting a multi-phase scheme similar to the method proposed in Vese and Chan (2002). h is the inverse edge image, where edges, or

more meaningfully the boundaries between two texture regions, are represented as lower intensity values. $\mathbf{F}^k_{in}$ and $\mathbf{F}^k_{out}$ are calculated iteratively, as the average PL histograms of inside and outside of the evolving contour.

Level set method implicitly define the evolving curve C by the zero level of the level set function $\phi(x, y)$ (Vese and Chan, 2001). $\phi(x, y)$ is simply the sign distance function (SDF) of C. Instead of $\phi(x, y)$, the shape characteristic function (SCF), $\chi(x, y)$, can also be employed as in Mahmoodi (2009). SCF is defined as,

$$\chi(x, y) = H(\phi(x, y)) \tag{A.2}$$

H is the Heaviside function. The shape-based active contour method to construct the evolution function has a faster convergence, less memory usage and better performance in the presence of noise Mahmoodi (2009). By replacing C in equation A.1 by χ, the terms in the energy functional E are expressed in following way. Let Ω be the image domain.

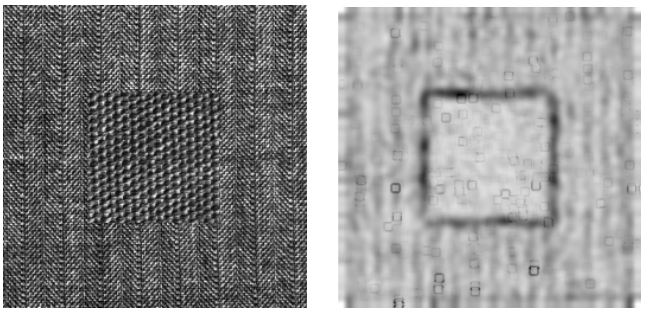
$$\begin{aligned} E(\mathbf{F}^k_{in}, \mathbf{F}^k_{out}, \chi) &= \mu \int_\Omega h(x, y)|\nabla\chi(x, y)|\,\mathrm{d}x\,\mathrm{d}y + \lambda_{in} \sum_{k=1}^{(n^2+1)/2} \int_\Omega D(\mathbf{F}^k_s, \mathbf{F}^k_{in})\chi(x, y)\,\mathrm{d}x\,\mathrm{d}y \\ &\quad + \lambda_{out} \sum_{k=1}^{(n^2+1)/2} \int_\Omega D(\mathbf{F}^k_s, \mathbf{F}^k_{out})(1 - \chi(x, y))\,\mathrm{d}x\,\mathrm{d}y \end{aligned} \tag{A.3}$$

The h is calculated via the metric of the 2D image manifold (Sagiv et al., 2006; Kimmel et al., 1997). Let the i^{th} bin image is represented by $\mathbf{W}^i = [W^i(x, y)|1 \leq x \leq Imwidth, 1 \leq y \leq Imheight]$. Then the metric of the feature space is given by,

$$g(x, y) = \begin{pmatrix} 1 + \sum_{i=1}^{L}(W^i_x)^2 & \sum_{i=1}^{L} W^i_x W^i_y \\ \sum_{i=1}^{L} W^i_x W^i_y & 1 + \sum_{i=1}^{L}(W^i_y)^2 \end{pmatrix} \tag{A.4}$$

And W^i_x and W^i_y are partial derivatives w.r.t. x and y. Partial derivatives can be obtained by finite difference method applied on to i^{th} bin image and L is equal to the number of total bins in the vector form concatenated PL histogram.

Since the metric g is associated with measuring distances on manifolds its components represent the rate of change of a given manifold in a specific direction. Hence the determinant of g has higher values when a strong gradient presents (Sagiv et al., 2006). Therefore metric g can be used for edge detection. Then the value $h(x, y)$ at (x, y) of inverse edge image h can be formed as,



(a) (b)

Figure A.1: (a)Texture image (b)Inverse edge image

$$h(x,y) = \frac{1}{det(g(x,y))} \tag{A.5}$$

where $det()$ represents taking determinant of a matrix and $g(x,y)$ represent matrix of g metric for the location (x,y). An inverse edge image obtained in this way for smoothed PL histogram features is show in figure A.1. In energy functional, given in equation A.1, the first term associates with the boundary information, extracted by inverse edge image, while the second and third terms relate to regional information. The evolution function obtained by applying Euler-Lagrange method for equation A.3 is,

$$\frac{\partial \chi}{\partial t} = \mu * div\left(\mathbf{h}\frac{\nabla \chi}{|\nabla \chi|}\right) - \lambda_{in}\sum_{k=1}^{(n^2+1)/2} D(\mathbf{F}^k, \mathbf{F}^k_{in}) + \lambda_{out}\sum_{k=1}^{(n^2+1)/2} D(\mathbf{F}^k, \mathbf{F}^k_{out}) \tag{A.6}$$

The shape based active contour segmentation algorithm proposed by Mahmoodi (2009) can be briefly stated as follows. The algorithm start with defining an arbitrary initial non smooth shape characteristic function (SCF) χ_0^0 for $t = 0$. The superscript and subscript indicate the iteration number and whether the regularization is applied or not. Next χ_0^0 is regularized to achieve χ_ϵ^0. Regularization is done by,

$$\chi_\epsilon^t = G_\epsilon(x,y) \star \chi_0^t \tag{A.7}$$

with,

$$G_\epsilon(x,y) = \frac{\epsilon}{\pi(x^2+y^2+\epsilon^2)} \tag{A.8}$$

where ϵ is the regularizing parameter. Then by equation (A.6) $\chi_{c\epsilon}^t$ at iteration t is calculated. To obtain the discrete version of equation (A.6) the semi implicit finite difference method discussed in Vese and Chan (2002) is used. Also at iteration t, χ_0^t is reconstructed by,

$$\chi_0^t = H(\chi_{c\epsilon}^t - \frac{1}{2}) \tag{A.9}$$

The regularized SCF for current iteration χ_ϵ^t is obtained by A.7. The regularized SCF χ_ϵ^t is normalized to obtain $\chi_{n\epsilon}^t$, which is between zero and one, and the zero crossings of $(\chi_{n\epsilon}^t - 1/2)$ is detected, which turns out to be the evolving contour at iteration t (Mahmoodi, 2009). The algorithm is illustrated as in algorithm A.0.1.

Algorithm A.0.1: *ShapeBasedActiveContourEvolution*(*FeaImgStack*)

Initialize χ_0^0 for $t = 0$ and regularize it and then use eq A.6 to calculate $\chi_{c\epsilon}^t$
repeat
Construct χ_ϵ^t by $\chi_{c\epsilon}^t$ using results of eq A.9 regularized by eq A.7 and normalize it to obtain $\chi_{n\epsilon}^t$
Detect zero crossing of $\left(\chi_{n\epsilon}^t - \frac{1}{2}\right)$ as the evolving contour at iteration t
Compute $\mathbf{F}_{in}^k$ and $\mathbf{F}_{out}^k$
Using eq A.6 to calculate $\chi_{c\epsilon}^{t+1}$ of next iteration
t=t+1
until convergence is reached ($D(\chi_0^t, \chi_0^{t+1}) < threshold$)

Appendix B

Gibbs Sampling Algorithm

The Gibbs sampling algorithm can be used to synthesize an image Y_{syn} when the conditional probability distribution of a pixel value given its neighbour values is known. The Gibbs sampling algorithm is given in Algorithm B.0.2 where L is the gray-levels and M is the number of sweeps (Li, 2009). In this study $L \in \{0, 1, \ldots, 255\}$ and $M = 200$ is used.

Algorithm B.0.2: GIBBS SAMPLER()

% Generating a texture using a Gibbs sampler.
begin
initialize Y_{syn} to white noise;
for $s \in S$ do
    compute $p_l = p(y_s = l | y_{s+r}, r \in N; \mathbf{f})$ for all $l \in L$;
    set y_s to l with probability p_l;
repeat for M times;
end

References

Ahonen, T. and Pietikäinen, M. (2008). A framework for analysing texture descriptors. In *3rd Int'l. Conf. on Computer Vision Theory and Applications*, pages 507–512.

Ahonen, T. and Pietikäinen, M. (2009). Image description using joint distribution of filter bank responses. *Pattern Recognition Letters*, 30(4):368–376.

Aykroyd, R. (1998). Bayesian estimation for homogeneous and inhomogeneous Gaussian random fields. *IEEE Trans. on pattern analysis and machine intelligence*, 20:533–539.

Besag, J. and Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, 82:733–746.

Bianconi, F. and Fernandez, A. (2007). Evaluation of the effects of Gabor filter parameters on texture classification. *Pattern Recognition*, 40:3325–3335.

Björkström, A. (2001). Ridge regression and inverse problems. Stockholm University, Department of Mathematics.

Blackmore, D., Danaher, L., and Radiopaedia.org (2014). Pulmonary emphysema. `http://radiopaedia.org/articles/pulmonary-emphysema`.

Blostein, D. and Ahuja, N. (1989). Shape from texture: Integrating texture-element extraction and surface estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(12):1233–1251.

Blunsden, S. (2004). Texture classification using non-parametric Markov random fields. *School of Informatics, University of Edinburgh.*

Blunsden, S. and Torrealba, V. (2005). Investigating the effects of scale in MRF texture classification. In *IET Int'l Conf. on Visual Information Engineering*, pages 275–280.

Bosnjak, A., Montilla, G., and Torrealba, V. (1998). Medical images segmentation using Gabor filters applied to echocardiographic images. In *Computers in Cardiology*, pages 457–460.

Bouman, C. A. (2009). MAP, EM, MRFs, and all that: A user's guide to the tools of model based image and signal processing. `https://engineering.purdue.edu/~bouman/ece641/previous/fall08/notes/pdf/Book-header.pdf`.

Brodatz, P. (1966). *Textures: A Photographic Album for Artists and Designers*. Dover.

Burt, P. J. and Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications*, 31(4):532–540.

Cesmeli, E. and Wang, D. (2001). Texture segmentation using Gaussian- Markov random fields and neural oscillator networks. *IEEE Trans. on Neural Networks*, 12(2):394–404.

Chellappa, R. and Chatterjee, S. (1985). Classification of textures using Gaussian Markov random fields. *IEEE Trans. on Acoustics Speech and Signal Processing*, 33(4):959–963.

Chen, Y. Q., Nixon, M. S., and Thomas, D. W. (1995). Statistical geometrical features for texture classification. *Pattern Recognition*, 28(4):537–552.

Comer, M. and Delp, E. (1999). Segmentation of textured images using a multiresolution Gaussian autoregressive model. *IEEE Trans. on Image Processing*, 8(3):408–420.

Coroyer, C., Declercq, D., and Duvaut, P. (1997). Texture classification using third order correlation tools. In *IEEE Signal Processing Workshop on High-Order Statistics*, pages 171–175.

Coxson, H. O., Rogers, R. M., Whittall, K. P., Yachkova, Y. D., Par, P. D., Sciurba, F. C., and Hogg, J. C. (1999). A quantification of the lung surface area in emphysema using computed tomography. *American journal of respiratory and critical care medicine*, 159 (3):851–6.

Dana, K., Nayar, Ginneken, S., B., and Koenderink, J. (1997). Reflectance and texture of real-world surfaces. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pages 151–157.

Data fact sheet (2003). Chronic obstructive pulmonary disease. `http://www.nhlbi.nih.gov`.

Davis, L. S. (1981). Polarograms: a new tool for image texture analysis. *Pattern Recognition*, 13(3):219–223.

Deng, H. and Clausi, D. A. (2004). Gaussian Markov random field rotation-invariant features for image classification. *IEEE Trans. on pattern analysis and machine intelligence*, 26(7):951–5.

Depeursinge, A., Iavindrasana, J., Hidki, A., Cohen, G., Geissbuhler, A., Platon, A., Poletti, P., and Müller, H. (2010). Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization. *Journal of digital imaging*, 23(1):18–30.

Depeursinge, A., Sage, D., Hidki, A., Platon, A., Poletti, P. A., Unser, M., and Muller, H. (2007). Lung tissue classification using wavelet frames. In *Proc. IEEE Int'l conf. of the Engineering in Medicine and Biology Society*, pages 6259–6262.

Descombes, X., Sigelle, M., and Preteux, F. (1999). Estimating Gaussian Markov random field parameters in a nonstationary framework: application to remote sensing imaging. *IEEE Trans. on Image Processing*, 8(4):490–503.

Dharmagunawardhana, C., Mahmoodi, S., Bennett, M., and Mahesan, N. (2012). Unsupervised texture segmentation using active contours and local distributions of Gaussian Markov random field parameters. In *Proc. of British Machine Vision Conference*, pages 88.1–88.11.

Duck, A. (2008). Nursing patients with interstitial lung disease. *Nursing Times*, 104(9):46–49.

Efros, A. and Leung, T. (1999). Texture synthesis by non-parametric sampling. In *Proc. of 7th Int'l Conf. on Computer Vision*, volume 2, pages 1033–1038.

Encyclopaedia Britannica (2006). Structure of lungs. `http://www.britannica.com/EBchecked/media/107200/`.

Encyclopaedia Britannica (2013). Emphysema. `http://www.britannica.com/EBchecked/media/107143`.

Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4):157168.

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer graphics and image processing*, 4(2):172–179.

Gangeh, M. J., Srensen, L., Shaker, S. B., Kamel, M. S., Bruijne, M., and Loog, M. (2010). A texton-based approach for the classification of lung parenchyma in CT images. *Medical image computing and computer-assisted intervention (MICCAI)*, 13:595–602.

Gomez, M. and Salinas, R. A. (2006). A new technique for texture classification using Markov random fields. *International Journal of Computers, Communications & Control*, 1:41–51.

Gonzalez, R. C., Woods, R. E., and Eddins, S. L. (2004). *Digital image processing using MATLAB*. Mc Graw Hill.

Greenspan, H., Belongie, S., Goodman, R., and Perona, P. (1994). Rotation invariant texture recognition using a steerable pyramid. 2:162–167.

Griffin, L. and Lillholm, M. (2010). Symmetry sensitivities of derivative of Gaussian filters. 32:1072–83.

Guo, Z. and Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Trans. on Image Processing*, 19(6):1657–1663.

Guo, Z., Zhang, L., and Zhang, D. (2010). Rotation invariant texture classification using lbp variance (lbpv) with global matching. *Pattern recognition*, 43(3):706–719.

Gupta, M. R., Garcia, E. K., and Chin, E. (2008). Adaptive local linear regression with application to printer color management. *IEEE Trans. on Image Processing*, 17(6):936–945.

Hadjidemetriou, E., Grossberg, M. D., and Nayar, S. K. (2003). Multiresolution histograms and their use for texture classification. In *Int'l Workshop on Texture Analysis and Synthesis*.

Haindl, M. and Mikeš, S. (2008). Texture segmentation benchmark. In *Proc. 19th Int'l Conf. Pattern Recognition*, pages 1–4.

Haley, G. M. and Manjunath, B. (1995). Rotation-invariant texture classification using modified Gabor filters. In *Int'l Conf. on Image Processing*, volume 1, pages 262–265.

Häme, Y., Angelini, E. D., Hoffman, E. A., Barr, R. G., and Laine, A. F. (2013). Robust quantification of pulmonary emphysema with a hidden Markov measure field model. In *Proc. 10th IEEE Int'l Symposium on Biomedical Imaging*, pages 382–385.

Hara, T., Yamamoto, A., Zhou, X., Iwano, S., Itoh, S., Fujita, H., and Ishigaki, T. (2004). Automated detection system for pulmonary emphysema on 3D chest CT images. In *Proc. of SPIE*, pages 915–919.

Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804.

Haralick, R. M., Shanmugam, K., and Dinstein, I. H. (1973). Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics*, (6):610–621.

Hastie, T., Tibshirani, R., and Friedman, J. (2013). *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. Springer-Verlag London Ltd, 2nd edition.

Hasudungan, A. (2014). Emphysema - Pathophysiology (COPD). `http://www.youtube.com/watch?v=ChlSfDBHLvg`.

Heeger, D. J. and Bergen, J. R. (1995). Pyramid-based texture analysis/synthesis. In *Proc. of 22nd conf. on Computer graphics and interactive techniques*, pages 229–238.

Heikkila, M. and Ahonen, T. (2012). Uniform Local Binary Patterns: Matlab code. `http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab`.

Held, K., Kops, E. R., Krause, B. J., Wells I., W. M., Kikinis, R., and Muller-Gartner, H. W. (1997). Markov random field segmentation of brain MR images. *IEEE Trans. on Medical Imaging*, 16:878–886.

Houhou, N., Thiran, J., and Bresson, X. (2009). Fast texture segmentation based on semi-local region descriptor and active contour. *Numerical Mathematics: Theory, Methods & Applications*, 2(4):445–468.

Huang, X., Dong, J., and Wang, M. (2011). Paper web defection segmentation using Gauss-Markov random field texture features. In *Proc. IEEE Int'l Conf. Image Analysis and Signal Processing*, pages 167–170.

Huang, Y. and Chan, K. L. (2004). Texture decomposition by harmonics extraction from higher order statistics. *IEEE Trans. on Image Processing*, 13(1):1–14.

Huang, Y., Chan, K. L., and Zhang, Z. (2003). Texture classification by multi-model feature integration using bayesian networks. *Pattern Recognition Letters*, 24:393401.

Jain, A. K. and Farrokhnia, F. (1990). Unsupervised texture segmentation using Gabor filters. In *Proc. IEEE Int'l Conf. Systems Man and Cybernetics*, pages 14–19.

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, 290(5802):91–97.

Karu, K., Jain, A. K., and Bolle, R. M. (1996). Is there any texture in the image? *Pattern Recognition*, 29(9):1437–1446.

Kashyap, R. L. and Khotanzad, A. (1986). A model-based method for rotation invariant texture classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, (4):472–481.

Kim, D. H., Yun, D., and Lee, S. U. (2006). New MRF parameter estimation technique for texture image segmentation using hierarchical GMRF model based on random spatial interaction and mean field theory. In *Proc. 18th Int'l Conf. Pattern Recognition*, pages 365–368.

Kim, N., Seo, J. B., Lee, Y., Lee, J. G., Kim, S. S., , and Kang, S. (2009). Development of an automatic classification system for differentiation of obstructive lung disease using HRCT. *Journal of digital imaging*, 22(2):136–148.

Kimmel, R., Malladi, R., and Sochen, N. (1997). Images as embedding maps and minimal surfaces: Movies, color, and volumetric medical images. *International Journal of Computer Vision*, 39(1):111–129.

Krishnamachari, S. and Chellappa, R. (1997). Multiresolution Gaussian–Markov random field models for texture segmentation. *IEEE Trans. on image processing*, 6(2):251–67.

Kwatra, V., Essa, I., Bobick, A., and Kwatra, N. (2005). Texture optimization for example-based synthesis. *ACM Transactions on Graphics*, 24:795.

Laws, K. I. (1980). Textured image segmentation. Technical report, DTIC.

Lazebnik, S., Schmid, C., and Ponce, J. (2005). A sparse texture representation using local affine regions. *IEEE Trans. on pattern analysis and machine intelligence*, 27(8):1265–1278.

Lee, C., Chen, S., Tsai, H., Chung, P., and Chiang, Y. (2006). Discrimination of liver diseases from ct images based on Gabor filters. In *19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06)*, pages 203–206.

Leung, T. and Malik, J. (2001). Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43(1):29–44.

Li, S. Z. (2009). M*arkov Random Field Modeling in Image Analysis*. Springer-Verlag London Ltd, 3rd edn edition.

Liao, S. and Chung, A. (2007). Texture classification by using advanced local binary patterns and spatial distribution of dominant patterns. In *IEEE Int'l Conf. on Acoustics, Speech and Signal Processing*, page 1221.

Litmanovich, D., Boiselle, P. M., and Bankier, A. (2009). CT of pulmonary emphysema - current status, challenges, and future directions. *European radiology*, 19:537–51.

Liu, X. and Wang, D. (2002). A spectral histogram model for texton modeling and texture discrimination. *Vision Research*, 42(23):2617–2634.

Liu, X. and Wang, D. (2003). Texture classification using spectral histograms. *IEEE Trans. on image processing*, 12:661–70.

Liu, X. and Wang, D. (2006). Image and texture segmentation using local spectral histograms. *IEEE Trans. on Image Processing*, 15(10):3066–3077.

Long, Z. and Younan, N. H. (2006). Contourlet spectral histogram for texture classification. In *IEEE Symposium on Image Analysis and Interpretation*, pages 31–35.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.

Lynch, D. A., Newell, J. D., and Lee, J.-S. (2000). *Imaging of Diffuse Lung Disease*. BC Decker Inc., 1st edition.

Madani, A., Keyzer, C., and AGevenois, P. (2001). Quantitative computed tomography assessment of lung structure and function in pulmonary emphysema. *The European respiratory journal*, 18:720–30.

Mäenpää, T. and Pietikäinen, M. (2005). Texture analysis with local binary patterns. *Handbook of Pattern Recognition and Computer Vision*, 3:197–216.

Mahmoodi, S. (2009). Shape-based active contours for fast video segmentation. *Signal Processing Letters*, 16(10):857–860.

Mahmoodi, S. and Gunn, S. (2011). Snake based unsupervised texture segmentation using Gaussian Markov random field models. In *Proc. 18th IEEE Int'l Conf. Image Processing*, pages 1–4.

Manjunath, B. S. and Chellappa, R. (1991). Unsupervised texture segmentation using Markov random field models. *IEEE Trans. on pattern analysis and machine intelligence*, 13(5):478–482.

Mao, J. and Jain, A. K. (1992). Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern recognition*, 25(2):173–188.

Martin, D., Fowlkes, Tal, D., and Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. 8th Int'l Conf. Computer Vision*, pages 416–423.

Matthews, T., Nixon, M. S., and Niranjan, M. (2013). Enriching texture analysis with semantic data. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 1248–1255.

Mayorga, M. A. and Ludeman, L. C. (1994). Shift and rotation invariant texture recognition with neural nets. In *IEEE Int'l Conf. Neural Networks*, volume 6, pages 4078–4083.

Mishima, M., T. Hirai, H. I., Nakano, Y., Sakai, H., Muro, S., Nishimura, K., Oku, Y., Chin, K., Ohi, M., Nakamura, T., Bates, J. H. T., Alencar, A. M., and Suki, B. L. (1999). Complexity of terminal airspace geometry assessed by lung computed tomography in normal subjects and patients with chronic obstructive pulmonary disease. In *Proc. the National Academy of Sciences*, volume 96, pages 8829–8834.

Movellan, J. R. (2002). Tutorial on Gabor Filters. `http://mplab.ucsd.edu/tutorials/gabor.pdf`. Open Source Document.

Muller, N., Staples, C., and R. Miller, R. A. (1988). Density mask, an objective method to quantitate emphysema using computed tomography. *Chest Journal*, 94:782–787.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47(1):90–100.

Nikou, C., Likas, C., and Galatsanos, N. P. (2010). A bayesian framework for image segmentation with spatially varying mixtures. *IEEE Trans. on Image Processing*, 19(9):2278–2289.

Nixon, M. and Aguado, A. (2008). *Feature extraction & image processing*. Academic Press, 2nd edition.

Office for National Statistics (2013). What are the top causes of death by age and gender? `http://www.ons.gov.uk/ons/index.html`.

Ohanian, P. P. and Dubes, R. C. (1992). Performance evaluation for four classes of textural features. *Pattern recognition*, 25(8):819–833.

Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59.

Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on pattern analysis and machine intelligence*, 24:971–987.

Ojala, T., Valkealahti, K., Oja, E., and Pietikäinen, M. (2001). Texture discrimination with multidimensional distributions of signed gray-level differences. *Pattern Recognition*, 34:727–739.

Olowoyeye, A., Tuceryan, M., and Fang, S. (2009). Medical volume segmentation using bank of Gabor filters. In *Proc. the symposium on Applied Computing*, pages 826–829.

Outex Texture Database (2007). Texture classification test suites, center for machine vision research, University of Oulu. `http://www.outex.oulu.fi/index.php?page=classification`.

Patil, S., Patil, H. S., et al. (2013). Study and review of various image texture classification methods. *International Journal of Computer Applications*, 75:33–38.

Peng, U., Wei, T. X., and Fu, F. J. (2005). A unified model of GMRF and MOG for image segmentation. In *Proc. IEEE Int'l Conf. Image Processing*, pages 1140–1143.

Penrose, R. (1956). On best approximate solutions of linear matrix equations. *Proc. Cambridge Philosophical Society*, 52:17–19.

Petrou, M. and Sevilla, P. G. (2006). *Image Processing, Dealing with Texture*. John Wiley & Sons Ltd.

Pichler, O., Teuner, A., and Hosticka, B. J. (1996). A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms. *Pattern Recognition*, 29(5):733–742.

Pietikäinen, M. (2010). Local Binary Patterns. `http://www.scholarpedia.org/article/Local_Binary_Patterns`. Scholarpedia.

Pietikäinen, M., Ojala, T., and Xu, Z. (2000). Rotation-invariant texture classification using feature distributions. *Pattern Recognition*, 33(1):43–52.

Porter, R. and Canagarajah, N. (1997). Robust rotation-invariant texture classification: wavelet, Gabor filter and GMRF based schemes. In *IEE Proc. Vision, Image and Signal Processing*, volume 144, pages 180–188.

Portilla, J., Strela, V., Wainwright, M. J., and Simoncelli, E. P. (2003). Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans. on Image Processing*, 12(11):1338–1351.

Recio R., J. A., Fernndez, L. A., and Sarri, A. (2005). Use of Gabor filters for texture classification of digital images. In *Fisica de la Tierra*, page 4759.

Rue, H. and Held, L. (2005). *Gaussian Markov Random Fields; Theory and Applications*. Chapman & Hall/CRC.

Sagiv, C., Sochen, N. A., and Zeevi, Y. (2006). Integrated active contours for texture segmentation. *IEEE Trans. on image processing*, 15(6):1633–46.

Simoncelli, E. P. and Freeman, W. T. (1995). The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Int'l Conf. on Image Processing*, volume 3, pages 3444–3444.

Singh, M. and Singh, S. (2002). Spatial texture analysis: a comparative study. In *Proc. 16th Int'l Conf. on Pattern Recognition*, volume 1, pages 676–679.

Sklansky, J. (1978). Image segmentation and feature extraction. *IEEE Trans. on Systems, Man and Cybernetics*, 8(4):237–247.

Sluimer, I. C., Waes, P. F., Viergever, M., and Ginneken, B. (2003). Computer-aided diagnosis in high resolution CT of the lungs. *Medical Physics*, 30:3081.

Sørensen, L., Shaker, S. B., and de Bruijne, M. (2010). Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans. on Medical Imaging*, 29(2):559–569.

Sørensen, L., Shaker, S. B., and de Bruijne, M. (2013). Computed Tomography Emphysema Database. `http://image.diku.dk/emphysema_database/`.

Sprawls, P. (1995). *The Physical Principles of Medical Imaging*. Medical Physics Publishing, USA, 2nd edition.

Srinivasan, G. and Shobha, G. (2008). Statistical texture analysis. *Proc. World Academy of Science: Engineering & Technology*, 36:1264–1269.

Stan, S., Palubinskas, G., and Datcu, M. (2002). Bayesian selection of the neighbourhood order for Gauss-Markov texture models. *Pattern recognition Letters*, 23:1229–1238.

Sullivan, B. and Liu, B. (1984). On the use of singular value decomposition and decimation in discrete-time band-limited signal extrapolation. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 32(6):1201–1212.

Swain, M. J. and Ballard, D. H. (1992). Indexing via color histograms. In *Active Perception and Robot Vision*, pages 261–273.

Takeda, H., Farsiu, S., and Milanfar, P. (2007). Kernel regression for image processing and reconstruction. *IEEE Trans. on Image Processing*, 16(2):349–366.

Tamura, H., Mori, S., and Yamawaki, T. (1978). Textural features corresponding to visual perception. *IEEE Trans. on Systems, Man and Cybernetics*, 8(6):460–473.

Topi, M., Matti, P., and Timo, O. (2000). Texture classification by multi-predicate local binary pattern operators. In *Proc. 15th Int'l Conf. on Pattern Recognition*, pages 939–942.

Tou, J. Y., Tay, Y. H., and Lau, P. Y. (2007). Gabor filters and grey-level co-occurrence matrices in texture classification. In *Int'l Symposium on Information and Communications Technologies*, pages 197–202.

Tou, J. Y., Tay, Y. H., and Lau, P. Y. (2009). Gabor filters as feature images for covariance matrix on texture classification problem. *Advances in Neuro-Information Processing*, 5507:745–751.

Tuceryan, M. and Jain, A. K. (1998). Texture analysis. In *Handbook of Pattern Recognition and Computer Vision*, chapter 2.1, pages 207–248. World Scientific Publishing Co., 2nd edition.

Uppaluri, R., Hoffman, E., M.Sonka, Hartley, P. G., W.Hunninghake, G., and McLennan, G. (1999). Computer recognition of regional lung disease patterns. *American journal of respiratory and critical care medicine*, 160 (2):648–54.

Valkealahti, K. and Oja, E. (1998). Reduced multidimensional co-occurrence histograms in texture classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20:90–94.

Varma, M. and Zisserman, A. (2002). Classifying images of materials: Achieving viewpoint and illumination independence. In *European Conf. Computer Vision*, volume 129, page 255271.

Varma, M. and Zisserman, A. (2005). A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81.

Varma, M. and Zisserman, A. (2009). A statistical approach to material classification using image patch exemplars. *IEEE Trans. on pattern analysis and machine intelligence*, 31:2032–2047.

Vasconcelos, V., Silva, J. S., Marques, L., and Barroso, J. (2010). Statistical textural features for classification of lung emphysema in CT images: A comparative study. In *Information Systems and Technologies*, pages 1–5.

Veni, S. (2010). Image enhancement of medical images using Gabor filter bank on hexagonal sampled grids. In *Engineering and Technology*, volume 2, pages 816–821.

Vese, L. and Chan, T. (2001). Active contours without edges. *IEEE transactions on image processing*, 10(2):266–277.

Vese, L. and Chan, T. (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293.

Waller, B., Nixon, M., and Carter, J. (2011). Texture segmentation by evidence gathering. In *Proc. 3rd BMVC Workshop*, pages 1–11.

Waller, B. M., Nixon, M. S., and Carter, J. N. (2012). Analysing micro-and macrostructures in textures. In *8th int'l conf. on Signal Image Technology and Internet Based Systems*, pages 246–253.

Wang, L. (1999). Texture classification using multiresolution Markov random field models. *Pattern Recognition Letters*, 20:171–182.

Webb, W. R., Muller, N. L., and Naidich, D. P. (2008). *High-resolution CT of the Lung*. Lippincott Williams and Wilkins, 4th edition.

Wikipedia (2014). Chronic obstructive pulmonary disease. `http://en.wikipedia.org/wiki/Chronic_obstructive_pulmonary_disease`.

Wikipedia (2014). X-ray computed tomography. `http://en.wikipedia.org/wiki/X-ray_computed_tomography`.

Xia, G., He, C., and Sun, H. (2006a). Urban extraction from SAR images using local statistical characteristics and Gaussian Markov random field model. In *Proc. 8th IEEE Int'l Conf. Signal Processing*, pages 20–23.

Xia, Y., Feng, D., and Zhao, R. (2006b). Adaptive segmentation of textured images by using the coupled Markov random field model. *IEEE Trans. on Image Processing*, 15(11):3559–3566.

Xie, X. and Mirmehdi, M. (2008). A galaxy of texture features. *Handbook of texture analysis*, pages 375–406.

Xu, Y., Sonka, M., McLennan, G., Guo, J., and Hoffman, E. A. (2006). MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. *IEEE Trans. on Medical imaging*, 156 (1):25.

Zhang, J., Marszałek, M., Lazebnik, S., and Schmid, C. (2007). Local features and kernels for classification of texture and object categories: A comprehensive study. *International journal of computer vision*, 73(2):213–238.

Zhang, J. and Tan, T. (2002). Brief review of invariant texture analysis methods. *Pattern recognition*, 35(3):735–747.

Zhang, J., Tan, T., and Ma, L. (2002). Invariant texture segmentation via circular Gabor filters. In *Proc. 16th Int'l Conf. Pattern Recognition*, pages 901–904.

Zhao, Y., Zhang, L., Li, P., and Huang, B. (2007). Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features. *IEEE Trans. on Geoscience and Remote Sensing*, 45(5):1458–1468.

Zhu, S., Zhang, L., and Jin, H. (2012). A locally linear regression model for boundary preserving regularization in stereo matching. In *European Conf. on Computer Vision*, pages 101–115.

Zhu, S.-C., Guo, C.-E., Wang, Y., and Xu, Z. (2005). What are textons? *International Journal of Computer Vision*, 62(1-2):121–143.

Zhu, S. C., Wu, Y., and Mumford, D. (1998). Filters, random fields and maximum entropy (frame): Towards a unified theory for texture modeling. *International Journal of Computer Vision*, 27(2):107–126.