Multimodal and multivariate texture representation based on the finite mixtures of generalized Gaussians with applications

par

Ndah Daniel Yapi

Thèse présentée au Département d'informatique et d'ingénierie en vue de l'obtention du grade de Docteur ès sciences (Ph.D.)

UNIVERSITÉ DU QUÉBEC EN OUTAOUAIS

Gatineau, Québec, Canada, 02 Novembre 2021

Ce mémoire a été évalué par un jury composé des personnes suivantes :

Professeur	Kamel Adi	Président du jury
Professeur	Mohand Said Allili	Directeur de recherche
Professeur	Halim Boutayeb	Membre interne du jury
Professeur	Abdelhamid Mammeri	Membre externe du jury

SOMMAIRE

La texture est une caractéristique fondamentale de nombreux types d'images, et sa représentation est l'un des problèmes essentiels et difficiles de la vision par ordinateur et de la reconnaissance de formes, qui reste un sujet de recherche intensive. L'enjeu principal est alors de développer des méthodes capables de résumer, représenter n'importe quelle texture, monochrome ou couleur dans une signature compacte qui incarne autant d'informations que possible permettant de distinguer cette texture d'une autre. Diverses méthodes de représentation texturale et de modélisation ont été proposées au cours des dernières décennies, mais le problème reste d'actualité. Récemment, des méthodes statistiques ont montré leur efficacité pour discriminer les textures, mais ces méthodes contiennent des imprécisions et des insuffisances concernant la représentation de certains types de textures tels que les textures couleurs dont les distributions des coefficients des transformées Muli-résolution exhibent une multi-variabilité et une multimodalité.

Notre objectif principal dans ce travail est de développer des modèles statistiques pour la représentation de texture qui abordent ces limitations pour une représentation de texture efficace. Premièrement, nous avons proposé un nouvel algorithme capable de détecter et de localiser les défauts dans la plupart des groupes de tissus. L'algorithme utilise l'apprentissage supervisé pour distinguer les tissus sans défaut des tissus défectueux, sur la base des signatures RCT-MoGG. En plus d'être compactes et rapides à calculer, ces signatures permettent une localisation précise des défauts. En fait, les défauts sont détectés lors de l'inspection de l'image en testant des patchs locaux à l'aide du classificateur bayésien

appris (BC). Notre approche peut traiter plusieurs types de textiles, des plus simples aux plus complexes. Des expériences sur la base de données TILDA ont démontré que notre méthode donne de meilleurs résultats par rapport aux méthodes de pointe récentes. Deuxièmement, nous avons généralisé le modèle, pour le cas des données multivariées et multimodales. Le nouveau modèle appelé MoMGG permet non seulement de modéliser la corrélation entre les coefficients avec un emplacement, des échelles et des orientations adjacents dans la même transformée multi-résolution, mais également entre des sous-bandes de différentes transformations (par exemple, des contours et des ondelettes) pour créer une signature de texture plus riche et plus représentative. La nécessité de développer un tel modèle unifié est de pouvoir représenter tout type de texture, qu'elle soit monochrome ou couleur. Nous avons appliqué notre approche pour la recherche d'image par le contenu en utilisant une divergence approximative de Kullback-Leibler (KLD) entre MoMGGs pour mesurer la similitude de texture. Les expériences menées sur certaines bases de données de référence ont démontré les capacités de nos modèles et leurs performances par rapport aux travaux récents de la littérature. De plus, des expérimentations sur la reconstruction d'images à partir du MoMGG ont montré de très bonnes performances, ce qui offre un potentiel énorme pour des applications telles que la super-résolution et la compression image/vidéo.

ABSTRACT

Texture is a fundamental characteristic of many types of images, and texture representation is one of the essential and challenging problems in computer vision and pattern recognition which remains a topic of intensive research. The main challenge is then to develop methods capable to summarize, represent any texture, monochrome or color in a compact signature that embodies as much information as possibles allowing distinguish this texture from another. Various methods of textural representation and modelling have been proposed in the recent decades, but the problem is still topical. Recently, statistical methods have shown their effectiveness to discriminate textures. But these methods contain imprecisions and insufficiencies concerning multimodal and multivariate texture representation.

Our main goal in this work, is to develop statistical models for texture representation that adress these limitations for effective texture representation. First, we have proposed a new algorithm that has capability to detect and locate defects in most of the fabric groups. The algorithm uses supervised learning to discriminate between defect-free fabrics from defective ones, based on RCT-MoGG signatures. In addition to being compact and fast to compute, these signatures enable accurate localization of defects. Actually, defects are detected during image inspection by testing local patches using the learned Bayesian classifier (BC). Our approach can deal with multiple types of textile fabrics, from simple to more complex ones. Experiments on the TILDA database have demonstrated that our method yields better results compared with recent state-of-the-art methods. Second, we generalized the model, for the case of multivariate and multimodal data. The new

model called MoMGG enables not only to model correlation between coefficients with adjacent location, scales and orientations within the same multi-resolution transform, but also between subbands of different transforms (e.g., contourlets and wavelets) to build richer and more representative texture signature. The need to develop such unified model is to be able to represent any type of texture, whether monochrome or color. We applied our approach for texture retrieval by using an approximate Kullback-Leibler divergence (KLD) between MoMGGs to measure texture similarity. Experiments on some benchmark databases. Obtained results have demonstrated the capabilities of our models and their performance compared to recent work in the literature. Furthermore, from our model, we carried out experiments on image reconstruction which showed very good performance, which offers enormous potential for applications such as super-resolution and image / video compression

REMERCIEMENTS

Tout dabord, je voudrais exprimer ma sincère gratitude à mon conseiller M. Said Allili pour le soutien continu, pour sa patience, sa disponibilité, sa motivation et ses immenses connaissances. Son précieux soutien et ses conseils m'ont aidé tout le temps de la recherche et de la rédaction de cette thèse. Je n'aurais pas pu imaginer avoir un meilleur conseiller et mentor pour mon doctorat. Je suis également extrêmement reconnaissant aux membres du jury, qui ont donné de leur temps précieux pour évaluer cette thèse. Je ne pourrais terminer sans remercier mon épouse qui m'a soutenu durant ces longues années de recherches qui ont conduit à cette thèse. Je dédie ce travail à mes enfants.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMES

D_{ACC} Accuracy

ACC_L Local Accuracy

AF Autocorrelation Function

ARR Average Retrieval Rate

BC Bayesian Classifier

CT Contourlet Transform

DFB Directional Filter bank

ERU Elementary repetitive unit

F_R false alarm rate

GD Geodesic Distance

GGD Generalized Gaussian

GGD Generalized Gaussian Distribution

GT Generalized t

HF High Frequency

PCA Principal Component Analysis

HPPCA Hierarchical Probabilistic PCA

ICA Independant Component analysis

KL Kullback-Leibler

KLD Kullback-Leibler Divergence

LBP Local Binary pattern

MCMC Markov Chain Monte Carlo

MCMC-KLD Markov chain Monte Carlo - Kullback

MHA Metropolis–Hastings algorithm

MG Multivariate Gaussian

MGG Multivariate Generalized Gaussian

MLD Multivariate Laplacian Distribution

MLE Maximum Likelihood Estimation

MML Mimimum Message Length

MoGG Mixture of Generalized Gaussian

MoMGG Mixture of Multivariate Generalized Gaussian

MoMGG+KLD+CCOS CCOS denotes cross-correlation

MoMGG+KLD+CD MoMGG based on daubechie 4 color-based descriptor

MoMGG+KLD+Sh Sh denotes image shift

MoMGG+KLD+CombT Transform combinations

MPE Multivariate Power Exponential

PCA Principal Component Analysis

PPCA Probabilistic Principal Component Analysis

RCT Redundant Contourlet Transform

RCT-MoGG Redundant Contourlet Transform - Mixture of Generalized Gaussian

RLP Redundant Laplacian Pyramid

SCT Standard Contourlet Transform

SDD Slope Difference Distribution

SIRV Spherical Invariant Random vectors

SSOCBS Small Scale Overcomplete Basis Set

Chapter 1

Introduction

1.1 Texture representation problem

Our visual world is richly filled with a great variety of textures, present in images, ranging from multispectral satellite views to microscopic pictures of tissue samples. As a powerful visual cue, texture provides a useful information in identifying objects or regions of interest in images. Texture is different from color in that it refers to the spatial organization of a set of basic elements or primitives (i.e., textons), the fundamental microstructures in natural images and the atoms of preattentive human visual perception [1] . A textured region will obey some statistical properties, exhibiting periodically repeated textons with some degree of variability in their appearance and relative position [2]. Textures may range from purely stochastic to perfectly regular and everything in between (see Fig. 1.1).

Figure 1.1: Texture images from the VisTex [3] databaseTexture

As a longstanding, fundamental and challenging problem in the fields of computer vision and pattern recognition, texture analysis has been a topic of intensive research since the last decades. This is due to its significance both in understanding how the texture perception process works in human vision as well as in the important role it plays in a wide variety of applications. Texture analysis covers a very wide spectrum of methods. It can be structured into three levels of concentrations : the low-level analysis which concerns the representation and modelling of texture, the middle-level analysis embraces several operations including discrimination, classification, segmentation, synthesis [4], and the high-level characterized by the applications derived from the middle-level such as remote sensing, medical imaging, inspection, detection of defects, content based image retrieval and so on [5, 6, 7, 8, 31]. Texture representation and modelling are basic steps for texture discrimination, classification and segmentation, so they represent a fundamental analysis operation. Texture representation, i.e., the extraction of features that describe texture information, is at the core of texture analysis. There are several methods to represent

textures, some of them are presented in the next section.

1.2 Texture representation methods

Various methods of textural representation and modelling have been proposed in recent decades, but the problem remains a subject of intensive research. Tuceryan and Jain [4] identified three major methods for texture representation: structural, spectral and statistical methods. Structural approaches represent texture by well-defined primitives (microtexture) and a hierarchy of spatial arrangements (macrotexture) of those primitives [1, 10]. Structural methods include Voronoi partitions [11] which characterize and segment textures by building a mosaic of Voronoi polygons and calculating descriptors on each Voronoi mesh. Topological invariant methods use for exemple texture descriptor which relies on Hough transform for finding line structures in highly structured textiles.

Spectral approaches represent an image in a space whose co-ordinates have an interpretation that is closely related to the characteristics of a texture (such as frequency or size). Among these methods, there are Fourier transform [12], Gabor filters [13, 14], wavelet transform [15], the Laplacian pyramids [16] , and contourlets transform [17]. The Fourier Transform allows to obtain a complex value module, called the Fourier spectrum which provides information on the power of the signal in each frequency. The other methods in this frequency group provide multiresolution representation allowing images to be successively decomposed, from coarse to fine resolutions. Wavelets and more recently contourlets have emerged as a signal processing effective tool to analyze texture information as they provide a multiresolution and orientation representation of an image. Among the most successful methods in texture representation and modelling there are statistical based methods [18].

In contrast to structural methods, statistical approaches do not attempt to understand

explicitly the hierarchical structure of the texture. Instead, they represent the texture indirectly by the non-deterministic properties that govern the distributions and relationships between the grey levels of an image. We can find the usual statistical reference method named cooccurrence matrix method, introduced by [19, 10], which provide spatial dependencies of the gray levels. Textures descriptors are generated from the co-occurrence matrices. Among statistical methods we also have random field method, Independent component analysis (ICA) [20], which is a factorial method which allows to explain a multidimentional signal in terms of components that are as independent as possible. We can also find many parametric statistical methods called generative methods or probabilistic methods or model-based models that describe the texture by a probabilistic model as an achievement of a random process [21]. Texture is then represented by the model parameters representing texture signature. Parametrics methods are the most powerful since they have proven their ability to successfully model texture [22]. Recently several parametric methods have been proposed, for example *probabilistic principal component analysis* densities, Weibull density, Generalized Gamma and Generalized Gaussian densities. A semi parametric modelling have been proposed using Generalized Gaussain density [23] which showed most flexibility and many image decompositions have been investigated jointly with. For these reasons, our method in this work focus mainly on generative models based on the generalized gaussian formulation.

1.3 Parametric models for texture representation

In the last decades, several parametric statistical models have been proposed to model multiscale transforms in texture image. The pioneering works used the parametric class of Gaussian scale mixtures, which allow to generate heavy-tailed distributions, by convolution of a normal distribution with a random one [21]. For example, generalized gamma, Student, Pearson Type VII, uniform power, generalized t (GT) and exponential power

distributions [21] have been used to model the distribution of wavelets. Among all these distributions, exponential distributions have been very successfully used for the representation and discrimination of multiscale transforms, in particular the generalized Gaussian density which has had a growing interest in texture modeling in the last decades. Recently, many works used the GGD model for gray level texture modelling in the frequency domain. The GGD as univariate distribution of wavelet subbands, has been applied in texture image retrieval and showed a good performance [23], [24], [7]. This success is achieved in part due to the fact that the GGD has a parameter that controls the shape of the distribution which provides some flexibility in the fitting of platykurtic and leptokurtic histograms of the transforms of most natural images.

Aside from ICA method, few researchers have attempted to address the multidimensional representation of texture. Studies have used multivariate generalized gaussian density (MGG) for representation of the distribution of spatial dependency in color texture [25], [26], [27]. Other models have been used such as spherical invariant random vectors (SIRV) [28] and models based on the copulas [29]. However, these multivariate models achieved an average performances, with sometimes too small number samples from generic texture database such as BRODATZ [34] et VISTEX [3]. Moreover, these methods have several restrictive assumptions that strongly limit the scope and comparability of their results, for example, most of the proposed models were all of zero mean as well as non-existent inter-band correlation. But it is well known that there is often a strong inter-band dependency at the frequency level. Also these works do not take into account the multimodalities observed in subbands for certain images, called multicomponent images [7]. This author [7] is one of the rare authors who modeled multimodality in the wavelet subbands, however this modeling was suitable for only with grayscale texture data i.e. univariates data, and not multivariate ones. The taking into account of mutimodility for multivariate data, for instance for wavelet subbands resulting from the decomposition of the three channels

(RGB) of color image, is still lacking.

It can be observed, through the aforementioned works, that the simultaneous modelling of multimodality and multidimensionality (in the sense of the correlation between data) in a unified study is still lacking. That is, previous works failed to address simultaneously, multicomponent and multidimensionality in texture representation. For example, to represent more accurately color texture taking account implicit correlations [25] observed between different subbands, it is necessary to use an appropriate full covariance model. Unfortunately this model does not exist to date, which is why we work on this in this work. In most of the cases, image representation methods extract the color features at each color channels independently, when in fact, there exist dependencies caused by linear transform in the color space and this fact has been highlighted in various works [25] . So, it is well known that there is often a strong inter-band dependency at the frequency level. Also, [22] highlighted other kind of correlations in grayscale image sub -bands, which capture the spatial distribution. The exploitation of these implicit inter-band, and spatial correlations in the multivariate framework could therefore lead to compact and robust signatures that embodies as much information as possible on the texture. It is this fact that motivated us to build-up a full covariance mixture of multivariate generalized Gaussian (MoMGG) model for texture representation.

More recently, deep learning methods using convolutional neural networks (CNN) have been proposed for the texture representation [30, 31, 32]. These methods learn filter weights and convolved labeled data with input images to create a spatial feature map of activations. They have the advantage of exploiting large datasets. However, they require the estimation of millions of parameters and a very large number of images which require enormous computing power, thus limiting their applicability with very little data. In the other and in [32, 33]., The authors found that the elaborated CNN models were not powerful for synthe-

sizing textures that humans cannot distinguish from natural textures. It would therefore be opportune to develop an efficient method, which could be used either upstream or downstream of such methods. A parametric method would be the best candidate to avoid these limitations.

Given previous works limitations, we need to develop an effective way that can improve and optimize the modelling and representation of the texture whether it is color or grayscale. In the context of applications such as Content-Based Image retrieval (CBIR), it is often necessary to discriminate a texture from others that have been described by models, and therefore it is necessary to develop also suitable similarity measures. There already exist similarity measures based on statistical distances. Some statistical methods have been used such as the Kullback Leibler distance (KLD) [23], Cauchy-Schwarz probability density function (pdf) distance [36, 37], the Rao geodesic distance (GD) [26] based on the Fisher information. In our work, we will propose appropriate approximations for the KLD in multivariate and multimodal in context of texture representation.

1.4 Contributions

Given the aforementioned challenges, we note that the main goal concerns improving precision for texture representing. The problem is to get a compact signature that is representative of the texture. First, this implies improving the existing methods of representation for univariate multicomponent textures, and then the build-up suitable models for multivalued, multicomponent textures. Herein, the approach considered is the parametric statistical characterization which, in addition to having demonstrated its ability to describe texture sparingly in reference works, can be used upstream and downstream of any texture representation method. With such advantages, statistical modeling still has a bright future ahead. Concerning multivalued multicomponent textures the problem is the lack of

unified model which fully exploit all the textural information contained in multivariate and multimodal subband distributions, for example for color texture. This latter challenge is addressed in one of the contributions of this work as a major contribution. Loosely speaking the general contribution of thesis, is to build-up a general framework which can handle not only unimodal or monochrome texture but also multivariate or multimodal texture. We can briefly summarize our main contributions in the following points:

1. In the first contribution, we propose an improved approach for applying MoGG model for local (in block sense) texture discrimination. Then, more specifically we applied the MoGG signatures in a blockwise discrimination leading to efficient texture classification. KLD distances between MoGGs were used as features which allows to discriminate defect from non-defective fabric texture. So far, MoGG has been most often used for texture retrieval. Here, we show that it is also valid as signature for texture discrimination. We adress this texture discrimination problem through a learning based approach using Bayes classifier. By opposite to several existing approaches for fabric defect detection, which effective in only some types of fabrics and/or defect, our method can deal with not only almost all type of pattern, that is motif-based and non motif-based one, but also all type of defects thanks to the RCT+ MoMG. The method enable both detection and localization of defects.

2. In a second contribution, we develop a new formal framework that extends the MoGG model to the mixture of multivariate generalized Gaussian (MoMGG) which allows representing and modelling both multimodality and correlations between multivariate data. Our method is an unified model for multivariate and multimodal distributions of coefficients in multi-resolution transforms of color texture images. The new MoMGG enables description of multivariate data with full covariance matrices. Strictly speaking, we propose a MoMGG with full covariance matrices to model arbitrary combinations of multi-scale decomposition bands. These can be bands originating from different color channels produced by different types of image transforms

(e.g., wavelets, contourlets, etc.) that many exhibit a strong inter-band and spatial correlations. So, the new MoMGG model enables not only to model correlation between coefficients with adjacent location, scales and orientations within the same multi-resolution transform, but also between subbands of different transforms (e.g., contourlets and wavelets) to build richer and more representative texture signature Given the flexibility of our MoMGG, we can fit any distribution of the combined coefficients without restrictions to previous work assumptions about unimodal and centered data. On the other hand, the MoMGG will enable accurate fitting of heavy-tailed mixture components which encourages more parsimonious models more disseminative capacity for texture classes. This second contribution is a major contribution in the sense that it paves the way for many applications such as discrimination of dynamic texture in video.

3. We then validate our MoMGG model in a context of content-based color image retrieval, where we used a multivariate approximation of Kullback-Leibler divergence (KLD) between MoMGGs using monte-carlo sampling in order to measure similarity between texture images.

1.5 Organisation of the thesis

Chapter 2 present a review of texture representaion and modelling methods. The principle of each method will be described and examples of their use will be presented. In chapter 3, since data of our models come from the spatio-frequency analysis notably multiresolution transform sub-bands, we will present the multiresolution transforms commonly used in texture analysis. Chapter 4 presents the generative methods used for textures representaion and modelling. Chapter 5 present our first contribution, the supervised Bayesian learning-base approache for Automatic Defect Detection in monochrome textured images. Chapter 6 presents the new framework of MoMGG parameters estimation and its application in

color texture retrieval as well as the multivariate approximation of the Kulback-leilber divergence. Through these applications we will show that our approaches far outweigh the works that has attempted to address the same topics. Finally, Chapter 7 concludes the work with a summary and discussion of the presented approaches and future work perspectives.

Chapter 2

A review of texture representation methods

2.1 Introduction

The term texture is commonly used to describe the structural components of images of natural scenes. In these images, we can observe more or less pronounced variations of intensity. Since the early work of Julesz [38], many methods of texture characterization have been proposed that rely on a range of approaches such as statistical analysis of spatial dependencies, structural analysis of textural patterns, frequency filtering of image content, probabilistic modelling. The goal sought in these numerous attempts to analyze the texture concerned mainly the extraction of sufficient distinguishing characteristics representing a compact signature of the texture. A compact signature is advantageously a large size reduction and allows automated processing of the texture. Unfortunately, there is no precise formal definition of texture, despite its omnipresence in almost all images (medical images, aerial, textiles, real scenes, etc.).

The pioneering work of Julesz [38] allowed to postulate that two textures are distin-

guished mainly by their first and second order statistics. This would mean that it is difficult to distinguish between two textures if their first and second order statistics are identical. [38] has also stated in his early work that it is difficult to distinguish between two textures if their three-order statistics are identical. Gagalowicz [39] has proposed counter-examples of textures having the same moments of orders three but which could easily be distinguished from each other. In the light of these observations, Julesz [1] established a new approach to interpreting the visual perception of textures, the texton theory. A texture in this approach would consist of primitives (texons). The human perception of the texture would then be governed by the first order statistics of the distribution of these primitives. The notion of primitives also provides a basis for the definition of texture made by [10] according to which a texture can be described by the number and type of its primitives as well as by the spatial organization or arrangement of the latter. Cross [40] argued that a texture is a stochastic two-dimensional field, possibly periodic. On the other hand, [41] and [42] have been among those who demonstrated the multi-frequency aspect of the visual perception of textures by the human cerebral cortex.

All these definitions have a common point: the notion of texture is used to translate a homogeneous aspect of the surface of an object into an image. The texture is manifested by a visual information that can describe it qualitatively according to the following criteria: coarse, fine, smooth, spotted, granular, mottled, regular or irregular etc. Texture can be seen as an important spatial feature, which can be used to identify objects or regions of interest in an image.

2.2 Texture representation methods

Texture is a ubiquitous cue in visual perception, and therefore an important topic within the science of vision. In particular, it has been studied in the fields of visual perception, computer vision and computer graphics [43, 44, 45]. All these fields require that texture

can in some way be mathematically modelled. Typically, a texture starts with a surface that exhibits local roughness or structure, which is then projected to form a textured image. Such an image exhibits certain aspects such as contrast, granularity, orientation, shape, regularity and roughness as the visual criteria that can be used for modelling. Due to the vague notion of texture, a multitude definitions have already been proposed in the literature. Generally, the analysis tends to be more driven by the desired application rather than any pure fundamentals.

This chapter introduces the most common methods in texture analysis. According to Tuceryan [46], these methods may be broken down into three broad categories of approaches : a structural approach, a statistical approach and frequency filtering approach. It should be noted that statistical approaches may belong to more than one of these categories. Also links can exist between descriptors resulting from methods of distinct categories.

2.2.1 Structural methods

The underlying assumption of structural methods is that textures consist of basic patterns called primitives repeating themselves in spatial directions in a quasi-regular way [10, 1]. Therefore, a structural description of a texture in this case involves the search for elementary patterns, their description, and then the determination of the rules of arrangement. These structural methods rely on an explicit definition of primitives and their placement rules. They are particularly well adapted to simple and regular textures said macroscopic. Voronoi partition according to [11] is one of the usual techniques of the category of structural methods. It is used to characterize and segment textures. The Voronoi partition consists of extracting the primitives texture (called tokens) and in building a mosaic of Voronoi polygons and calculating descriptors on each Voronoi mesh using the moments of the surfaces of the polygons. Another method is a topological invariant method [47] which shows a topologically invariant texture descriptor which relies on Hough transforms for finding line structures in highly structured textiles.

Another recent structural approach often classified as a statistical approach is *local binary partterns* (LBP) [48]. The original LBP texture operator is a rotation-invariant measure of a gray level texture, computed from pixel neighborhood analysis. LBP operator labels the pixels of an image with decimal numbers, called LBP codes, which encodes the local structure around each pixel. It proceeds thus, as illustrated in Fig. 2.1,

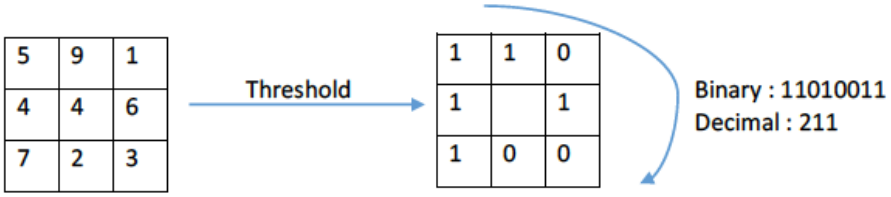


Figure 2.1: An example of the original LBP operator

where each pixel is compared with its neighbors in a 3x3 or 7x7 neighborhood by subtracting to them the center pixel value. The results that are strictly negative are put to 0 and the others to 1. A binary number is obtained by concatenating all these binary codes in a clockwise direction starting from the top-left one and its corresponding decimal value is used for labeling. The derived binary numbers are referred to as LBP codes. The method is considered to be mixed, combining the structural and statistical approach for texture analysis. The LBP method has many variants. For example, to deal with the texture at different scales, the operator was later generalized to use neighborhoods of different sizes. For more readings consult [49, 50], a detailed presentation can be seen in [51].

2.2.2 Spectral methods

Julesz [1] has highlighted the ability of the human brain to analyze images according to their frequency content. This caused the proliferation of texture analysis methods using a

frequency filtering of the image. They are motivated by the assimilation of perception to a selective, multi-scale and multi-orientations filtering mechanism [1]. Different methods of texture analysis (structural, statistical or model-based) can be exploited in a transformed space. Frequency approaches that have been used in recent decades for texture analysis are: The Fourier transform [12], Gabor filters [13, 14], the wavelet transform [15], the Laplacian pyramids [16], and contourlets transform [17]. Apart from the Fourier transform, the other four transform methods are used as multiresolution techniques in texture analysis. Multiresolution representation allows images to be successively approximated, from coarse to fine resolutions. Multiscale and multiresolution analysis maximizes the simultaneous localization of energy in both spatial and frequency domains. These tools are those that motivated the signal processing methods such as Gabor filters and wavelet transform. Wavelet, contourlet transforms and Laplacian pyramids will be developed in more depth in subsequent chapters due to the fact that they will be the basis of our contributions, i.e. our new approaches developed in this thesis.

2.2.2.1 The Fourier transform

The Fourier transform allows, in the domain of the image, to obtain a spatial frequency representation of the data. This contains a real part and an imaginary part, and we are mainly interested in its module, called the Fourier spectrum (See Fig. 2.2). Concerning texture, the Fourier transform highlights the frequency characteristics of a texture. Indeed, the spectrum allows accounting for the energy distribution of the image, to respect the periodicity as well as the orientation of the patterns of the texture image, which is particularly useful in the study of textures. Although the Fourier spectrum provides information on the location of frequencies, it does not allow any temporal and spatial location in the spectrum. It is therefore not very suitable for the detection and localization of textures.

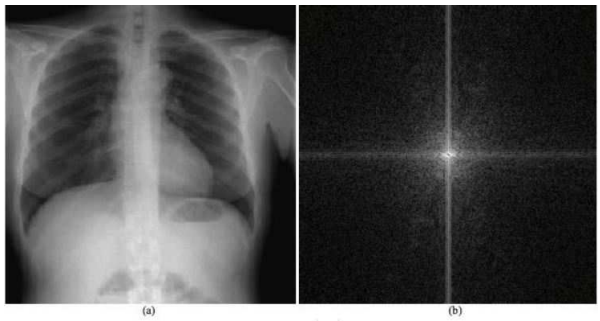


Figure 2.2: A chest radiograph is illustrated in (a), its 2-D Fourier spectrum in (b)

2.2.2.2 Gabor filters

The Gabor Filters have received considerable attention for texture representations because the characteristics of certain cells in the visual cortex of some mammals can be approximated by these filters. In addition, these filters have been shown to possess optimal localization properties in both spatial and frequency domain, and thus are well suited for many application in textural analysis [52, 53, 54, 55]. A Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope. A typical configuration of the use of Gabor filters in texture analysis is to use a filter bank as illustrated in Fig. 2.3. Each filter corresponds to a channel decomposing the image in certain frequency bands and orientation. We are talking about multichannel filtering. Low-pass filters are used in image denoising, high-pass filters in edge detection. They allow a good localization, both in the frequency and space domains. One limitation of Gabor filter is that it is high dimensional and computational intensive.

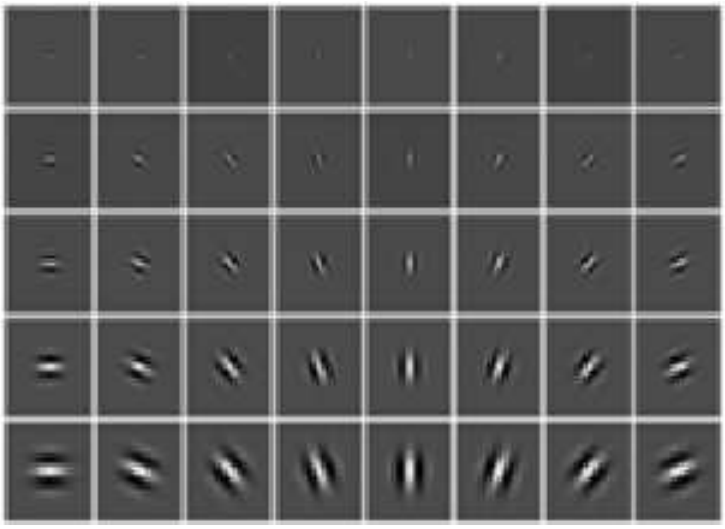


Figure 2.3: Fourty Gabor filters used to extract the Gabor features: Each row represent a different scale (wavelengths top to down: 3, 6, 12, 24, 48); whereas each column stands for a different orientation.

2.2.2.3 The Laplacian pyramid

The Laplacian Pyramid (non-oriented) has been used for multiresolution synthesis of textures [16]. It has been also used by [17, 56] for texture classification. Initially introduced by Morlet [57] for the analysis of seismic signal, it gives access to a multi-resolution spectral representation of the initial image. It is a powerful tool for spatio-frequency localization. The Laplacian pyramid was applied by [103] to image compression to remove redundancy. Compared to the Gaussian pyramid, the Laplacian pyramid is a more compact representation. Each level of a Laplacian pyramid contains the difference between a low-pass filtered version and an upsampled "predication" from coarser level, e.g.:

$$\mathbf{I}_L^{(n)} = \mathbf{I}_G^{(n)} - S \uparrow \mathbf{I}_G^{(n+1)}, \tag{2.1}$$

where $\mathbf{I}_L^{(n)}$ denotes the nth level in a Laplacian pyramid, $\mathbf{I}_G^{(n)}$ denotes the nth level in a Gaussian pyramid of the same image, and $S \uparrow$ represents upsampling using nearest neighbours.

2.2.2.4 The Wavelet transform

Wavelet transform, unlike the Gabor method, provides a uniform multiscale analysis requiring a single parameterization for all scales and allows the image to be decomposed into orthogonal and independent subbands, thereby, limiting the information redundancy. It is has been adopted by many authors, especially in segmentation [58, 59, 60], classification [24, 61, 62] or in content-base image retrieval [63, 23, 64]. However, the wavelets are limited in two points: directional selectivity and translational invariance [65]. The three possible analysis directions limit the detection capability of certain contours by the wavelets. Therefore, several extensions have been proposed to take into account oriented textures or presenting singular geometrical features not taken into account by the classical transforms. Among these extensions, we mention the bandelets [66], contourlets [67], the ridgelets [68] and curvelets [69].

2.2.2.5 The Contourlet transform

The Contourlet transform (CT) introduced by N. Do et al. [70, 71] is one of a multiresolution methods that recent texture analysis research has been focused on. The CT bears to the major limit of the two-dimensional discrete wavelet transform. It allows capturing more directional information in the image by a process of multiresolution and multidirectional decomposition of the image by iterations of filter banks using non-separable quincunx filters. It performs a multiresolution directional image representation that can efficiently capture and represent singularities along smooth objects in natural images. The decomposition is therefore carried out in two stages which produce low-frequency (LF) and high-frequency (HF) coefficients from the original input image as shown in the CT general structure illustrated in Fig. 2.4.

The first stage uses a Laplacian pyramid to transform the image into a series of band-pass levels and a low-pass level (low-frequency approximation of the image). The second stage appropriately applies $2D$ directional filter bank (DFB) and a critical sampling to

decompose each LP bandpass level into a number of directional bands, thereby capturing directional or oriented information. The DFB application is called directional subband decomposition (DSD). The DFB decomposition at each pyramidal level is defined by a vector, $nLevs$.

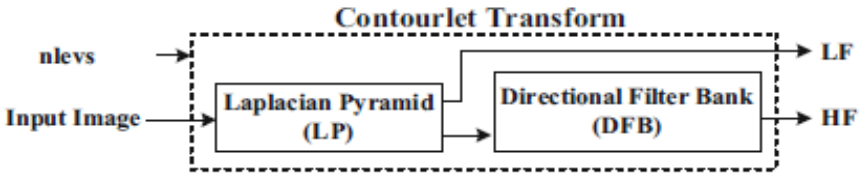


Figure 2.4: Process of the Contourlet Transforms

2.2.3 Statistical methods

The past decades have seen a strong emergence of statistical methods for texture analysis and representation. In these methods, texture is described by a collection of statistics of selected features. They are based on Julesz's finding [38, 72, 1] that the human visual system uses statistics features for texture discrimination because visual perception of texture is sensitive to their statistical properties. Some statistical methods are based on spatial analysis considering the spatial dependence of pixels, others are based only on order statistics while others consider texture as random process that can be describe by a probabilistic law. In spatial analysis, methods consist of studying the relationships between each pixel and his neighbors. They follow from the first conjectures of Julesz [38] in evaluating the distribution of the gray levels contained in the texture by statistical order descriptors computed according to a distance d and a particular orientation θ between a central pixel and its neighboring pixels. Among the methods in this category, one finds the matrices of co-occurrence which constitutes the reference in texture analysis and often serves as comparative method for the new approaches.

2.2.3.1 Cooccurrence matrix method

The cooccurrence matrices introduced by [19, 10] constitute a two-order method representing the spatial dependencies of the gray levels. They allow determining the frequency of observation of a pattern formed of a pair of pixels to which a translation is applied (which is decomposed in a direction θ and a translation d). The performance of this method is dependent on the choice of translation. Several statistical parameters serving as descriptors of the texture are then generated from the co-occurrence matrices. The most used of these descriptors are *angular second moment*, *variance*, *energy*, *contrast*, *entropy*, and *correlation* which allow describing the global aspect of the image. The problem with this approach is in that using grey-scale images discards information contained in the differences of hue and saturation that may provide further relevant information for classification tasks, thereby making the co-occurrence method less accurate. On the other hand, co-occurrence matrices do not provide any measure of texture that can easily be used for classification. The information in the matrix needs to be extracted as a set of values, such as entropy, correlation or homogeneity.

Among the statistical methods allowing to describe the global aspect of the image, we can find order statistics, which measure the probability of occurrence of a gray level at a random location without considering the spatial dependence of pixels. These statistics are calculated from the probability distribution (histogram) of the light intensity. The most commonly used study parameters are mean, variance, flattening (skewness), dissymmetry (kurtosis) and signal-to-noise ratio, which reflects the heterogeneity of the texture.

2.2.3.2 Grey-level run length method

The Grey-level run length method, a higher order method has been also used for texture analysis. A run is called a set of consecutive pixels in a given direction with the same gray level. The length of a run is then the number of pixels for a given run. The texture

descriptor in this method is then a set of matrices. Each matrix element (i, j) contains the number of runs of length j and gray level i in the direction θ. The most important statistical attributes calculated in this method are, the number of beach lengths, the portion of small and large beaches as well as the heterogeneity of the run lengths. One will be able to consult [73, 74, 75] for more details on these methods .

2.2.3.3 Independent component analysis (ICA)

Recently, new methods qualified as statistical methods have emerged. These do not proceed to the study of the relations between each pixel and its neighborhood as it is the case with the traditional statistical methods. One of the most famous is the Independent Component Analysis (ICA). The ICA method processes vectorial (multivariate) observations to extract linear components that are "as independent as possible". In other words, it seeks to explain a vector n-dimensional (which can represent a signal) in terms of components that are linear and statistically independent. Under the assumption that each of the components is one-dimensional and that the number of components is equal to the dimension of the observation, the model can then be written in the following form:

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{2.2}$$

where $\mathbf{A}$ is mixing matrix $n \times m$ invertible and $\mathbf{s}$ vector $m \times 1$ whose coefficients are statistically independent. Originally, this method was introduced to separate blindly. The different independent sources that compose a mixed signal. Blindly means without having any prior information on the sources. In the image processing domain, the signal to be processed is two-dimensional image. We start on the basis that each window of the image is considered as a linear combination of components that are statistically independent [20].

2.2.3.4 Multiresolution auto-regressive model

In statistics, econometrics and signal processing, an autoregressive (AR) model is a representation of a type of a random process; as such, it is used to describe certain time-varying processes in nature. An autoregressive model specifies that the output variable depends linearly on its own previous values and on a stochastic term; thus the model is in the form of a stochastic difference equation. Comer [76] has proposed the multiresolution autoregressive model for the characterization of texture in a multi-scale representation using a Gaussian pyramid. [77] and [78] have developed new grayscale and color texture segmentation methods, respectively based on AR models. Dubuisson et al [79] have proposed a characterization of textures by simultaneous autoregressive model (SAR), then segmented the image using maximum likelihood. In [80] textures are also characterized by the SAR model but the segmentation is done using markov random field models.

2.2.3.5 Parametric models

Generative or probabilistic methods, also called model-based methods, describe the texture probabilistically as an achievement of a random process that can be described by a probability law. Texture is then represented by the model parameters constituting a sort of texture signature. Among these model we can cite Hidden marckov chain, Markov random field, etc.

One of the oldest approaches in this statistical parametric models category is the generalized gaussain dstribution (GGD) [15] which is a family that includes the Gaussian and Laplacian distributions, to fit an observed histogram. Basically, the GGD model is a function of two parameters, namely, scale and shape parameters. The scale parameter represents the standard deviation of the probability density function while the shape parameter controls the shape of the GGD. The GGD has received a lot of interest and is widely used in different areas.

Recently several generative models have been proposed for modeling wavelet coefficients using distribution using parametric distribution families [81, 82, 83, 84]. Among these we can find product Bernoulli distribution (PBD) and its associated signature so-called bit-plane probability (BPP) signature [85, 86], generalized Gamma density [87], hierarchical probabilistic principal component analysis [88]. However another recently popular methods so-called mixture of Generalized Gaussian MoGG [89, 7], [90] has been developed and have obtained the best success so far. All of these methods are only dealing with monochrome textures.

2.2.4 Conclusion

In this chapter, we have outlined a non-exhaustive set of methods for texture representation. We started by exploring the meaning of texture. Typically, a texture represents a surface that exhibits local roughness or structure, which is then projected to form a textured image. Such an image exhibits both regularity, randomness to varying degrees on the one hand, and directionality and orientation on the other hand. These features have been relevant in texture representation and modelling. Techniques that have been used for this purpose include co-occurrence, Markov random fields, hidden Markov models, auto-correlation, run length method, ICA, Fourier transform, Gabor filters, GGD, and so on. These techniques can be structured into three broad categories of approaches : structural approaches, statistical approaches and frequency filtering approaches. Almost all of the above texture representation methods have limitations in term of precision, computational efficiency, invariance to translation, to rotation, robustness to noise, etc. Statistical parametric models have shown encouraging performance, particulary the generalized Gaussian densities in frequency domain.

Among the succesfull methods for texture representaion, multiresolution transforms

have shown a good performance. The reason is that the level of detail within an image varies from location to location. Finer resolution for analysis is required at regions where significant information is contained. Multiresolution representation of an image provides complete detail about the extent of information present at different locations. The main concept of multiresolution analysis is that for each vector space, there is another vector space of higher resolution until the final image is obtained. The basis of each of these vector spaces is the scale function. For textures, it provides scale invariant interpretation of a texture. Different multiresolution analysis tools will be present in next chapter.

Chapter 3

Texture representation using multiresolution transforms

3.1 Overview

Generally structures to be analyzed in an image have different sizes, which eliminates the existence of a single resolution adapted to all the objects present in the image. Therefore, it is desirable to analyze these structures in a progressive way using multiresolution representation which allows interpreting the image content in a hierarchical way. Objects at different resolutions characterize then the different physical structures present in the scene. Multiresolution analysis can be interpreted as the vision of the same image successively with a microscope, a magnifying glass and a human eye moving away progressively [15]. A multiresolution analysis considers that any signal can be built iteratively by adding details at different scales. The concept of multiresolution analysis introduced by Mallat [15] comes from the pyramidal algorithms of Burt and Adelson [91]. In the case of a digital image, these algorithms generate successive and increasingly coarse approximations of the original image. Multiresolution analysis is a special case of pyramidal algorithms. In this chap-

ter, we focus on the multiresolution analysis commonly used in texture analysis: Gabor, wavelets, standard and redundant contourlets.

3.2 Multichannel Gabor filters

Gabor filters are used to model spatial summation properties of simple cells in the visual cortex and have been adapted and popularly used for texture analysis [92, 93, 94]. They have been long considered as one of the most effective filtering techniques to extract useful texture features at different orientations and scales. Gabor filters can be categorized into two components: a real part as the symmetric component and an imaginary part as the asymmetric component. The 2D multichannel Gabor filter is a windowed signal processing method. It can be used for texture analysis for several reasons: they have tunable orientation and radial frequency bandwidths, tunable center frequencies and optimally achieve joint resolution in spatial and frequency domains. A very detailed analysis of the Gabor function using localized spatial filters for texture feature extraction is given in [95, 96]. A typical 2D Garbor model used in the texture analysis is given as follows:

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right] exp(2\pi j u_0 x + 2\pi j v_0 y) \tag{3.1}$$

where σ_x and σ_y define the Gaussian envelope along the x and y directions respectively, u_0 denotes the radial frequency of the Gabor function, and $j = \sqrt{-1}$. Fig. 3.1 shows the frequency response of the dyadic Gabor filter bank with the centre frequencies $2^{-\frac{11}{2}}$, $2^{-\frac{0}{2}}$, $2^{-\frac{7}{2}}$, $2^{-\frac{5}{2}}$, $2^{-\frac{3}{2}}$, and orientation 0^0, 45^0, 90^0, 135^0 [97].

Gabor filters provide a mean for better spatial localization. However, their usefulness is limited in practice because there is usually no single filter resolution at which one can localize a spatial structure in natural textures

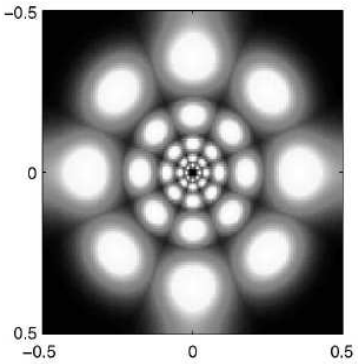


Figure 3.1: The frequency response of the dyadic bank of Gabor filters. The maximum amplitude response over all filters is plotted. Each filter is represented by one centre-symmetric pair of lobes. The axes are in normalised spatial frequencies.[97]

3.3 Wavelets transforms

Wavelet based texture analysis uses a class of functions that are localised in both spatial and frequency domain to decompose texture images. It is a very powerful model for texture desscription. The wavelet transform decomposes a texture image into a set of frequency channels that have narrower bandwidths in the lower frequency regions. The transform is suitable for textures consisting primarily of smooth components, so that their information is concentrated in the low frequency region.

A wavelet is a small wave that has a beginning and an end. Contrarly to the the Fourier transform which performs signal decomposition in a sinusoidal summation, the wavelet transform decomposes the input signal (finite energy function) into a weighted sum of small wavelets (wavelets). These elementary functions, also called analytic functions, are generated by translation and/or expansion of a unique function ψ, called wavelet-mother. Below are the analytic functions (wavelets) in the one-dimensional case:

$$\psi_{a,b} = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{3.2}$$

where a and b represent respectively the scale factor which plays the role of the inverse of a frequency and the translation factor which determines the position. The wavelet decomposition of an image $f(x, y)$ is then the convolution product between the image $f(x, y)$ and the wavelet functions $\psi_{a,b}$:

$$C_{(a,b)} = \int_{-\infty}^{+\infty} f(x, y) \frac{1}{\sqrt{ab}} \psi \left(\frac{x-b}{a}, \frac{y-c}{a} \right) dxdy \tag{3.3}$$

The wavelet coefficients therefore contain information about the image at different scales. The wavelets can then be considered as a filter bank composed of bandpass filters with relatively constant bandwidths. Moreover, the wavelet ψ and its associated scale function sometimes have no analytical form and therefore the use of their expressions in a filter form is preferable. Fig.3.2 presents an example of generation of a wavelet basis. The result of the decomposition is a set of coefficients called wavelet coefficients C from which we can also reconstruct the original function, i.e. the image $f(x, y)$. The functions or wavelet basis oscillate at a frequency equal to $1/a$ around the vector (b, c) creating a kind of "zooming" at the neighborhood of the considered point. Thus, the wavelet transform decreases its scale in high frequency and increases it in low frequency. The wavelet coefficients therefore contain information concerning the image at different scales.

3.3.0.1 Implementation of the wavelet transform

We focus in this work on the 2D discrete wavelet transform (DWT). It allows image representation in the spatio-frequency domain and it is based on multiresolution analysis which hierarchically breaks down a two-part signal. The first part, low frequency, is a kind of average of the original signal commonly called approximation image. The second part is a set of subbands in the image details; at each level of resolution, details are organized as three oriented subbands *horizontally*, *verticaly* and *diagonally*. These details are commonly referred to as wavelet coefficients. The applications of wavelet transforms to two-dimensional signals dates from Mallat [15]. He is at the origin of the fast decomposition/reconstruction

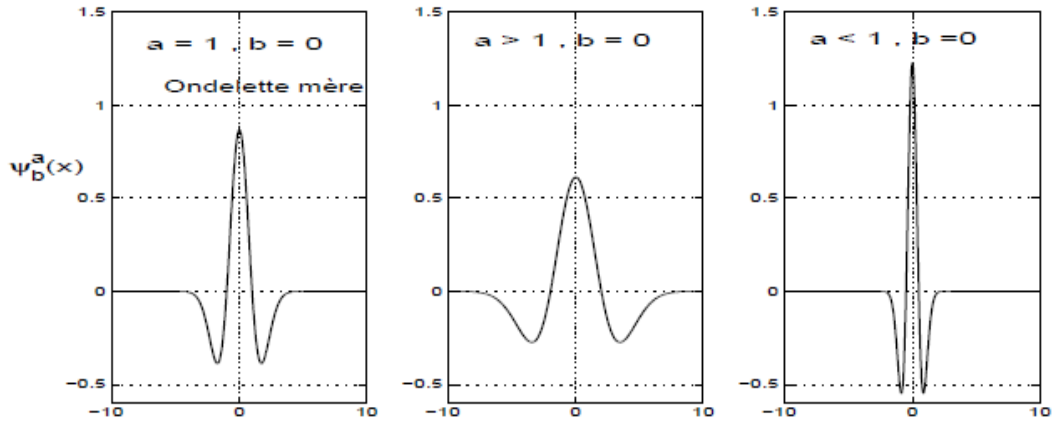


Figure 3.2: Examples of the "Mexican hat" wavelet family for some values of a. Analytical equation $\psi(x) = \left(\frac{2}{\sqrt{3}}\pi^{-1/4}\right)(1-x^2)exp(-x^2/2)$.

algorithm by wavelets. This algorithm is recursive and essentially based on two operations :

- **Filtering** : convolution of the original signal with a low-pass filter (h_0) or a high-pass filter (g_0).
- **Downsampling**: reducing the number of samples of the signal. Indeed, horizontal downsampling ($1:2$) leads to eliminate one column out of two, which reduces the number of pixels per line of half.

Fig.3.3 shows the diagram of Mallat's algorithm which can be explained as follows: Let S_j approximation image at a resolution level j and let D_j^X be the orientation subband X extracted at resolution level j. As shows in the figure, the image S_j goes into the input of the algorithm. It undergoes both filtering *highpass* and *lowpass*. The two resulting images are subsampled on the lines. The two subsampled images are each filtered by a *highpass* and a *lowpass* filters to produce four images. The latter are subsampled again giving thus 4 images of the same size: an approximation image (S_{j+1}) and three details images $D_{j+1}^X, X \in \{H, V, D\}$.

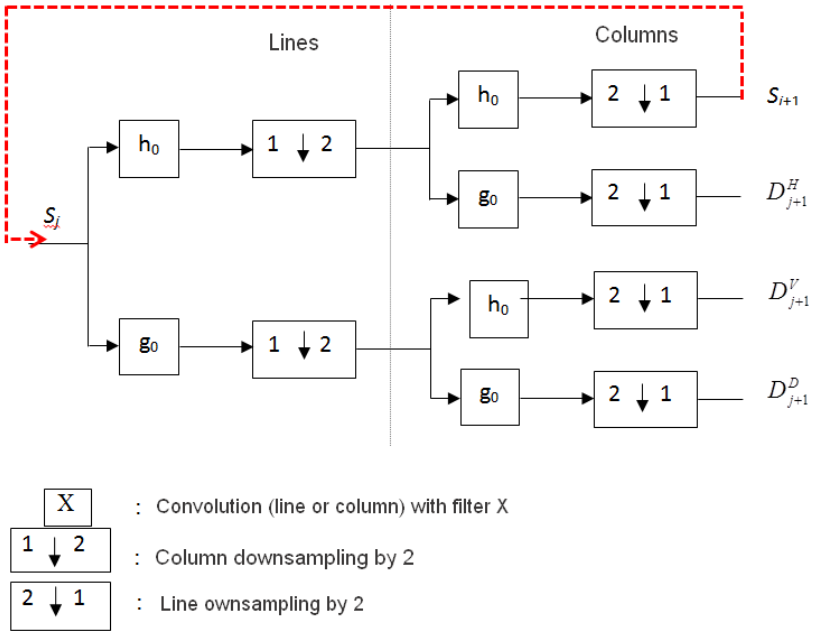
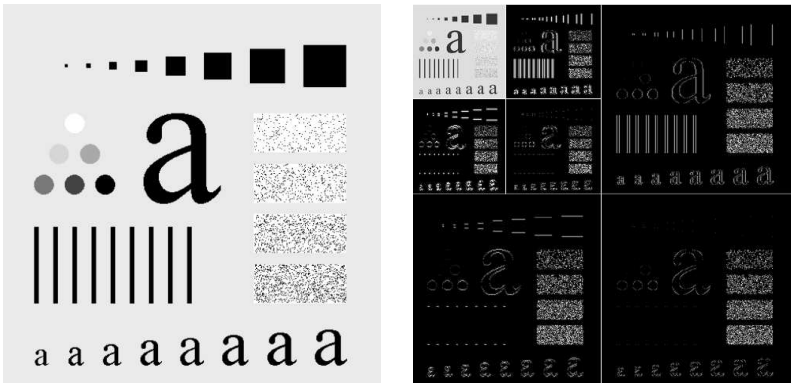


Figure 3.3: Recursive wavelets decomposition diagram and obtaining the four subbands.

Indeed, the scanning (filtering) of the lines leads to decomposing the image vertically into left approximation coefficients and into straight detail coefficients. After each scan of the rows, the column scan allows to decompose the image horizontally in two parts approximation and detail. The image is thus analyzed in different directions; the horizontal details are contained in D^H, vertical details in D^V and diagonal details in D^D. Block S contains an image of lower resolution (4 times less pixels than the initial image) whose details in different directions have been subtracted (principle of multiresolution). Each of these sub-images contains some of the information of the original image. Blocks D^H, D^V and D^D contain details of the image, or high-frequency information. In the case of decomposition at more than one level, the decomposition procedure is iterated on the approximation subband S. Fig.3.4 shows an example of the application of the discrete wavelet transform of a pattern with two decomposition levels.



(a) (b)

Figure 3.4: Example of wavelet decomposition. (a) original image, (b) Result of the image decomposition by DWT with two decomposition levels.

In addition to benefiting from a fast decomposition/reconstruction algorithm, the wavelet transform offers several other advantages. It offers a compact representation in that the total number of wavelet coefficients is equal to the number of pixels of the associated image. Wavelet coefficients are real coefficients in contrast to the Fourier transform (FFT) which provides complex coefficients. Note also that the reconstruction of the image from its wavelet representation is perfect and this shows that all the information contained in the image is preserved in its wavelet transform.

Several families of discrete wavelets can be used. These offer various forms of compromise between the regularity of the function and the more or less compact aspect of its support. Among these families, Daubechies filters [98] remain one of the most used wavelet base for texture analysis because of their quasi-fractal properties, the wide variety of functions they allow and the orthogonality of the subbands produced favoring perfect reconstruction. The different filters of this family are defined by their number of vanishing moments. We thus speak of Daubechies filters 2 (db2), 4 (db4), 6 (db6), 8 (db8), etc., functions going from quasi-fractal to softer ones with a growing number of zero moments. The simplest function, the Haar wavelet, which represents the special case of Daubechies filters has one vanishing moment. Haar wavelet uses a door function as scale function.

Chang and kuo [99] have used the tree-structured wavelet transform for texture classification. [100, 101] develops the wavelet transform for invariant texture analysis based on the Daubechies four-tap wavelet filter coefficients. In his method, the texture is decomposed into 10 channels, which are obtained through the three level wavelet decomposition. In [102], wavelet transform is used both to analyse the image prior to segmentation enabling feature selection as well as to provide spatial frequency-based descriptors for segmenting textures. The quality and accuracy of segmentation ultimately depend on the type of features used. Images consisting of a number of textured regions are best segmented using frequency-based features, whereas images made up of smoother regions can more easily be segmented using local mean and variance of intensity levels.

Wavelets have a limited ability in the representation of the directional information of the image, which causes a weakness in smooth contours detection. This limit comes from the limited number of orientations captured by wavelet filters during decomposition. It only allows three directions of analysis, horizontal, vertical or diagonal. This weakness led to the search for new forms of two-dimensional discrete transforms. Among the most recent and relevant, we can cite the contourlet transforms and their redundant variant [70, 71] .

3.4 The contourlet transform

The contourlet transform (CT) introduced in [70, 71] bearing to the major limit of the two-dimensional discrete wavelet transform. It consists of two parts, laplacian pyramid (LP) and directionnal filter bank (DFB). LP comprises a decimation part and an interpolation part. These two parts produce low-frequency (LF) and high-frequency (HF) coefficients from the original input image [103]. DFB analyzes the high-frequency coefficients, generated by LP, to show directionality in the frequency domain. CT can be constructed hierarchically applying low-frequency coefficients (LF) from one LP (e.g., $n1$ in Fig. 3.5(a)) to the input of another LP (e.g., $n2$ in Fig. 3.5(a)) as illustrated in Fig. 3.5(a).

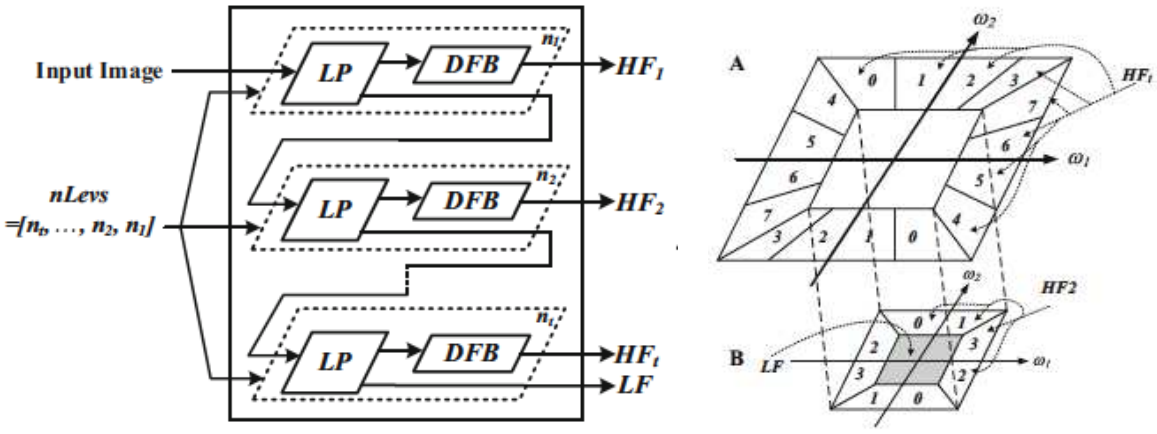


Figure 3.5: hierarchical block diagram of CT (a), 2D frequency plane decomposition for nLevs = [2, 3] (b)

The DFB decomposition at each pyramidal level is defined by a vector, $nLevs = [n_t \ldots n_2 n_1]$. Fig. 3.5(b) illustrates the 2D frequency plane decomposition of nLevs = [2, 3]. It shows two sets of directional sub-bands (A and B) that are eight and four wedgeshaped directional sub-bands.

3.4.1 The Laplacian pyramid

The construction of the LP is based on the manipulation of a Gaussian pyrzamid whose main purpose is to represent an image at decreasing levels of resolution [91]. The LP is the dual element of the Gaussian pyramid. More precisely, a level G_k of the LP is the result of a subtraction of a level G_k of a Gaussian pyramid and its estimated version by expansion of the upper level $G_k + 1$. Each level of the Laplacian pyramid represents an estimation error, due to the expansion algorithm used.

To generate the LP of an image G_0 low-pass filtering followed by a downsampling according to each dimension to obtain a low-pass filtered image G_1 and whose total size

is reduced by 4 compared to the original image G_0. The iteration of this filtering process gives the Gaussian pyramid. The difference (prediction error) between a level G_k of the Gaussian pyramid and its version estimated by expansion (oversampling with interpolation filter) of the upper level, leads to a level L_k of the LP. L_k therefore represents an estimation residue associated with the k level resolution. The initial prediction error image denoted L_0 is obtained by the following equation:

$$L_0 = G_0(i,j) - R_0, \tag{3.4}$$

where R_0 is an interpolated version of G_1 which have the same size as the image S_0 and g is a low-pass filter. Iteration of this process leads to a sequence of prediction error images L_0, L_1, L_2,...,L_n such that the size of each image is the quarter-size of its predecessor. This sequence is the LP of the image G_0. Thus, the decomposition stores the high frequencies in the bottom stages of the LP as can be seen in the example of the Fig. 3.6.

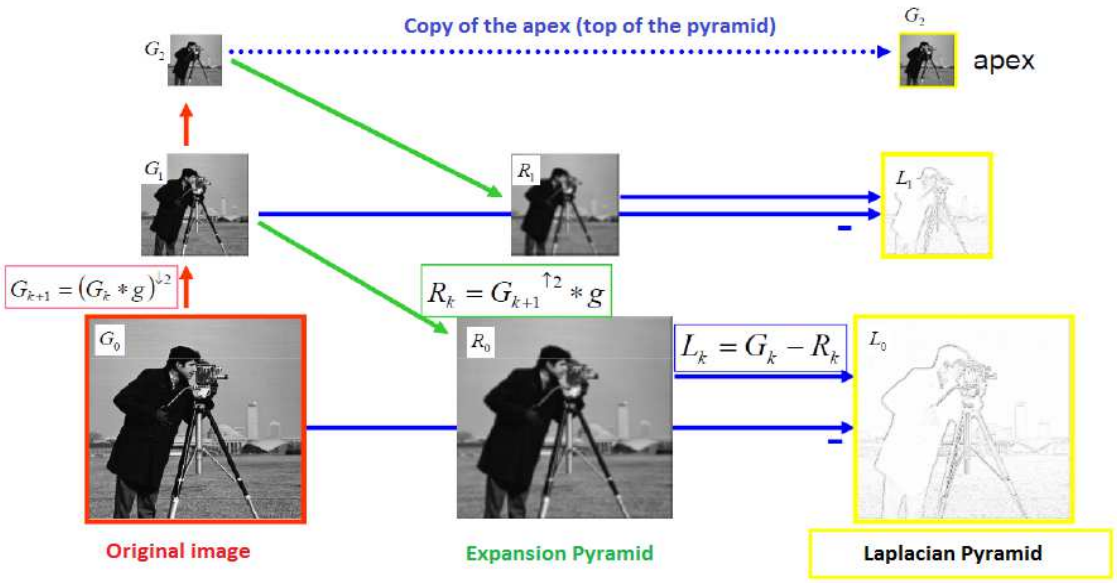


Figure 3.6: An example of laplacian pyramid decomposition

Indeed, L_k is widely decorrelated and its binary representation requires fewer bits than

the image G_k. In addition, the image G_k is filtered low-pass and can then be encoded at a reduced sampling rate. The Gaussian pyramid, which is the basis of the LP decomposition, is computed as follows. The original image is convolved with a Gaussian kernel whose values of the discrete coefficients approximate the values of a Gaussian curve. As described above, the resulting image is a low-pass filtered version of the original image. The Laplacian is then computed as the difference between the original image and the low pass filtered image. This process is continued to obtain a set of band-pass filtered images (since each is the difference between two levels of the Gaussian pyramid). Thus the LP is a set of band pass filters. The kernel typically used is 5×5 and the coefficients of this kernel must meet certain constraints relating to their separability, their normalization, their symmetry and their equicontribution [91]. Fig. 3.7 shows the generation process of the approximation image and the prediction error.

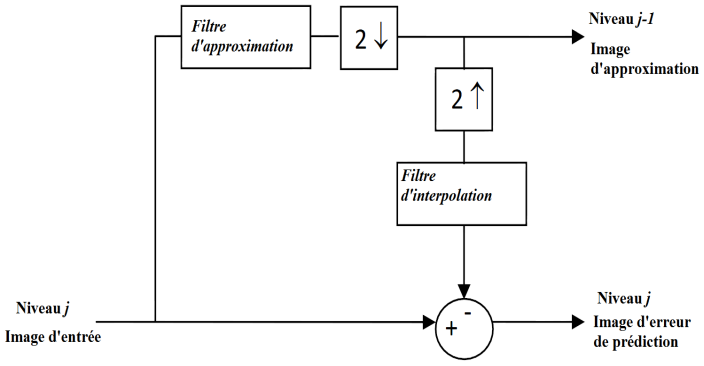


Figure 3.7: Laplacian Pyramid decomposition procedure: generation of the approximation and the prediction error.

3.4.2 Directional filters bank (DFB)

The original $2D$ directional filter bank (DFB) constructed by Bamberger and Smith [104] that uses two-dimensional quincunx filter banks with diamond-shaped filters and a critical decimation, decomposes the image into some components containing the directional informations.

The DFB is efficiently implemented via an l-level binary tree decomposition that leads to $2l$ subbands with wedge-shaped frequency partitioning as shown in Fig. 3.8.
For the contourlet transform, it is a simplified version of the original DFB that avoids modulating the input image and has a simpler rule for expanding the decomposition tree, which is used in the decomposition process. The simplified DFB used for the contourlette transformation consists of two steps and leads to subbands with frequency partitioning in conical form [105], where l is the decomposition level.

The first stage of the DFB is a two-channel quincunx filter bank (see Fig. 3.9.) [106] with fan filters that divide the $2D$ spectrum into vertical and horizontal directions, while the second stage is a shearing operator, which reorders the image samples. By adding a pair of $45 - degree$ shearing operator and its inverse (unshearing), before and after a two-channel filter bank (Fig. 3.9.), a different directional frequency partition is obtained (diagonal directions), while maintaining the ability to perfectly reconstruct the original image . Thus, the key in the DFB is to use an appropriate combination of shearing operators together with two-direction partition of quincunx filter banks at each node in a binary tree-structured filter bank, to obtain the desired 2-D spectrum division as shown in Fig. 3.8.

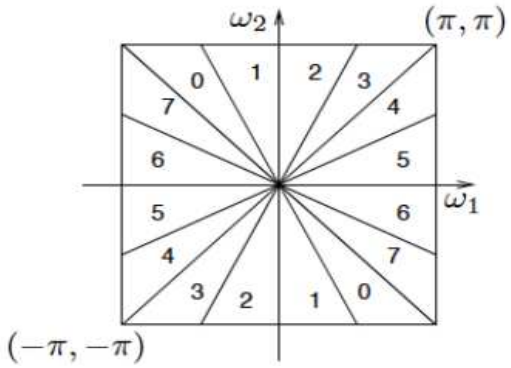


Figure 3.8: Directional filter bank. (a) Frequency partitioning where $l = 3$ and there are $2^3 = 8$ real wedge-shaped frequency bands. Subbands $0 - 3$ corresponds to the mostly horizontal directions, while subbands $4 - 7$ corresponds to the mostly vertical directions

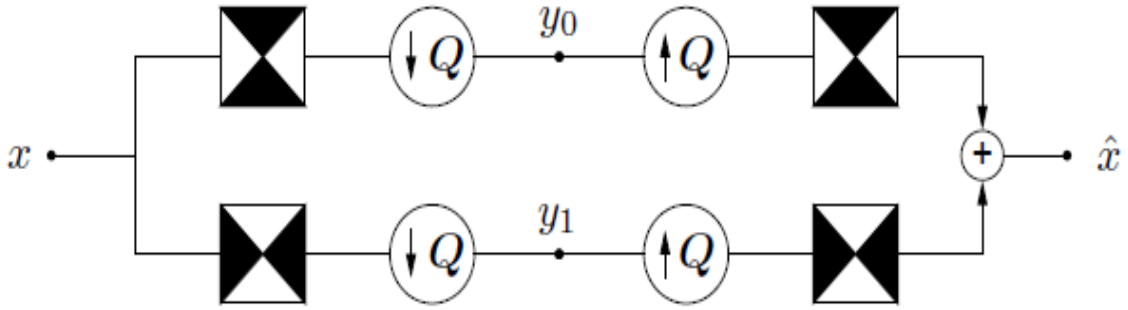


Figure 3.9: Two-dimensional spectrum partition using quincunx filter banks with fan filters. The black regions represent the ideal frequency supports of each filter. Q is a quincunx sampling matrix.

3.4.3 The standard contourlet transform

The combination of the LP and the DFB is a dual filter bank called *Pyramidal Directional Filter Bank* (PDFB) or contourlet filter bank (see Fig. 3.10). In order to capture directional information, bandpass images from LP decomposition are fed into a DFB so that directional information can be captured. This scheme can be repeated on the levels of approximation images. The combined result is the contourlet filter bank, which is a double iterated filter bank, which decomposes the images into directional subbands at multiple scales. In addition to offering perfect reconstruction and high computational efficiency, the CT is

almost critically sampled with a redundancy factor up to 4 : 3 due to the Laplacian pyramid. When compared to the DWT, the CT yields some improvements and new potentials in texture analysis applications with its improved directional selectivity [107]. To better understand the benefit of contourlets, we present in Fig. 3.11 an example of application of the wavelet transform and the contour line transform on an image, thus illustrating the operation of these two transforms. Part (a) of the examples of five images based on 2-D wavelets. Part (b) shows examples of four images based on contour lines. In part (c), we present an illustration showing how wavelets with square supports cannot capture the points of discontinuities, while contourlets with elongated supports can capture linear segments of contours, and thus can effectively represent a smooth contour with less coefficients [67]. Contourlets offer a much richer set of directions and shapes, and thus they are more effective in capturing smooth contours and geometric structures in images.

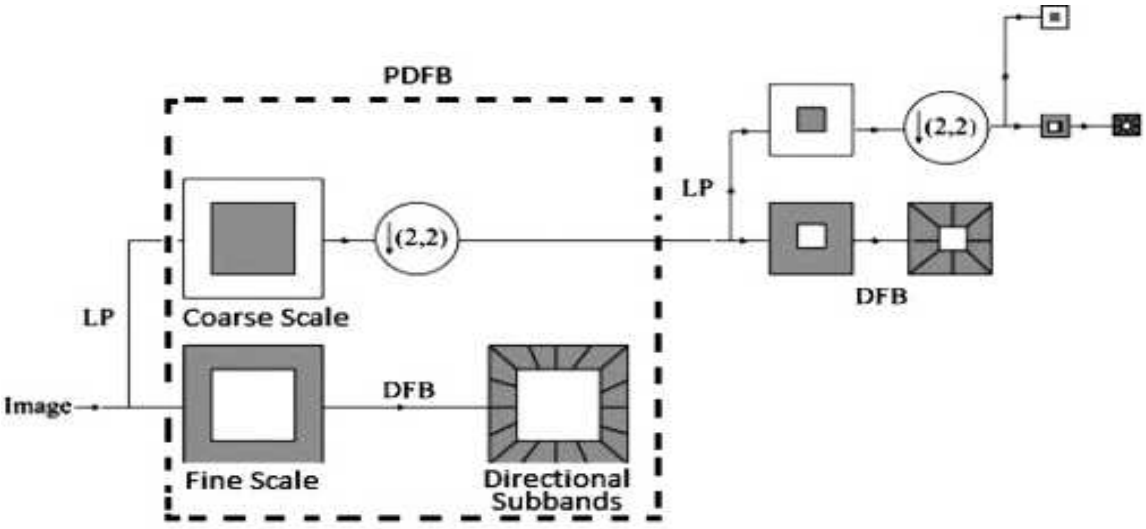


Figure 3.10: Contourlet filter bank.

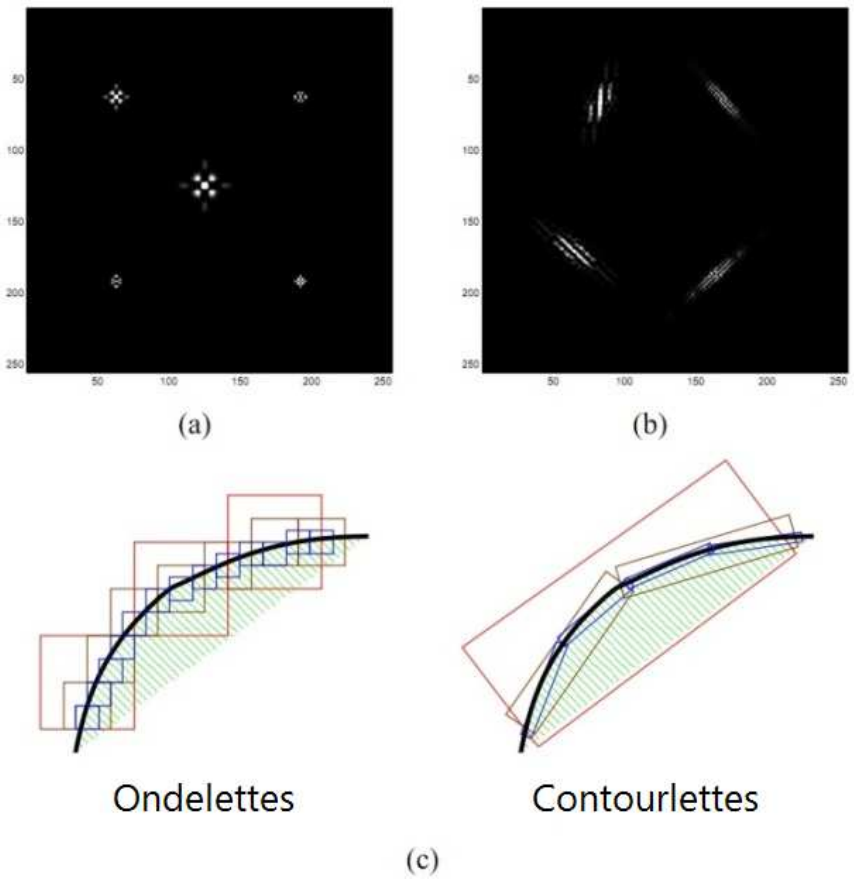
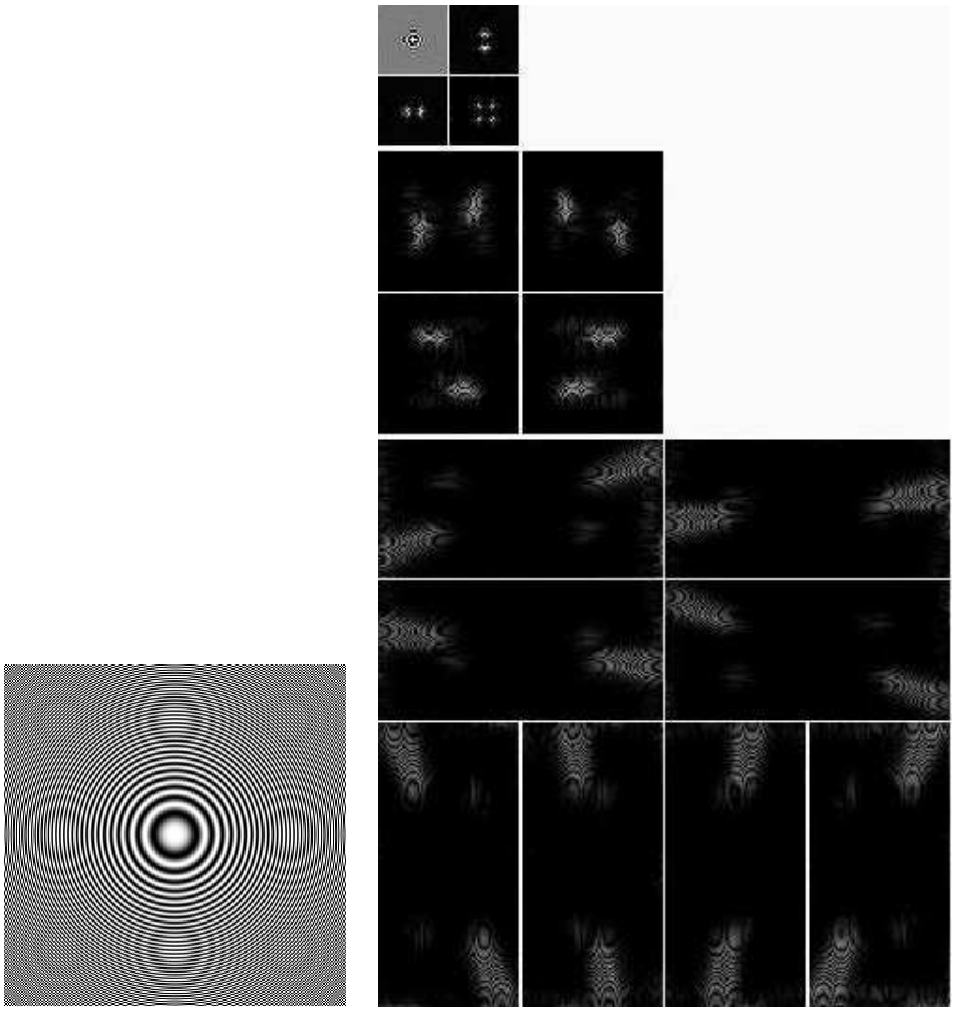


Figure 3.11: Comparing wavelet Vs. contourlet: on the left five $2D$ wavelets and four $2D$ contourlets on the right.

In Fig. 3.12, we show the execution of the contour transformation algorithm on an image called 'zoneplate'. The image passed as an input parameter of the contourlet transform algorithm is represented by (a) while (b) represents its contourlet transform. At first glance, it is clear that this is a 2-level contour contour transform ($L = 2$). We also notice that the first level admits 4 frequency directions while the second level admits 8 frequency directions. In this example, the Laplacian pyramid has 2 levels and 1 approximation. The bank of contourlets has subdivided L_0 into 8 sub-bands (2 levels). L_1 is subdivided into 4 sub-bands (1 level) and the approximation image is subdivided into 4 sub-bands (also 1 level). Fig. 3.12 exhibits clearly orientations in all directions other than vertical, horizontal and diagonal.



(a) Zoneplat image (b) 2-level contourlet decomposition

Figure 3.12: Example of application of the 2-levels contourlet transform

3.4.4 The Redundant contourlet transform (RCT)

There are a few variants of the discrete contourlet decomposition according to the specific applications. One of the most important is the redundant contourlet transform, which provides a major modification consisting to maintain a high redundancy in the contourlet subbands [108]. This change ensures an image representation scheme where all directional subbands are equal size (having the same number of contourlet coefficients), facilitating and increasing the precision of treatments that appeal to cooperation or the fusion of several

subbands. In general, in image analysis, hierarchical processes that require cooperation or mapping between different levels of resolution are made easier when all these resolution levels are the same size (in number of samples). The RCT shares the same decomposition scheme with the SCT. However, all downsampling operations in the RCT are discarded from the Laplacian stage achieving the redundancy of the contourlets transform.

By using L low-pass filters with the appropriate frequency characteristics, the L low-pass approximations of the original image are created. The difference between each approximation and its subsequent low pass version is a bandpass image. The final result is a redundant Laplacian pyramid (RLP) with $L + 1$ levels of equal sizes; an approximation image and L bandpass images. The redundancy factor is equal L given that each resolution level has the size of the original image and since the pyramid has L + 1 levels. Applying the same directional decomposition on the D dirirections desired (by a DFB with pseudo Gaussian filters and critical sampling) on each redundant Laplacian level, we obtain a redundant contourlet decomposition (RCD) with $L * D$ directional subbands of same sizes, plus the low-pass approximation of the image having the same size as the original image. Fig. 3.13 illustrates the multiscale representation and directional partition of the RLP and RCT in the frequency domain and Fig. 3.14 shows the whole diagram of the RCT. At the end of the process, each Contourlet subband is represented by a subimage $C_{l,d}$ where l is its resolution level and d its frequency direction (or frequency orientation).

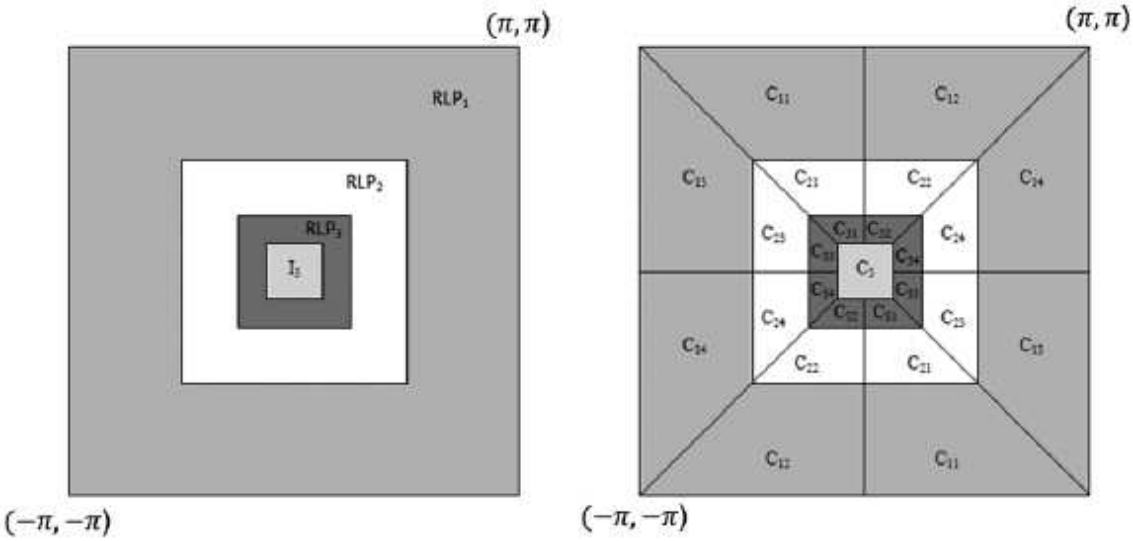


Figure 3.13: Three level RCT decomposition scheme ($L = 3$) and corresponding frequency partition. No downsampling operations performed at the Laplacian stage [108].

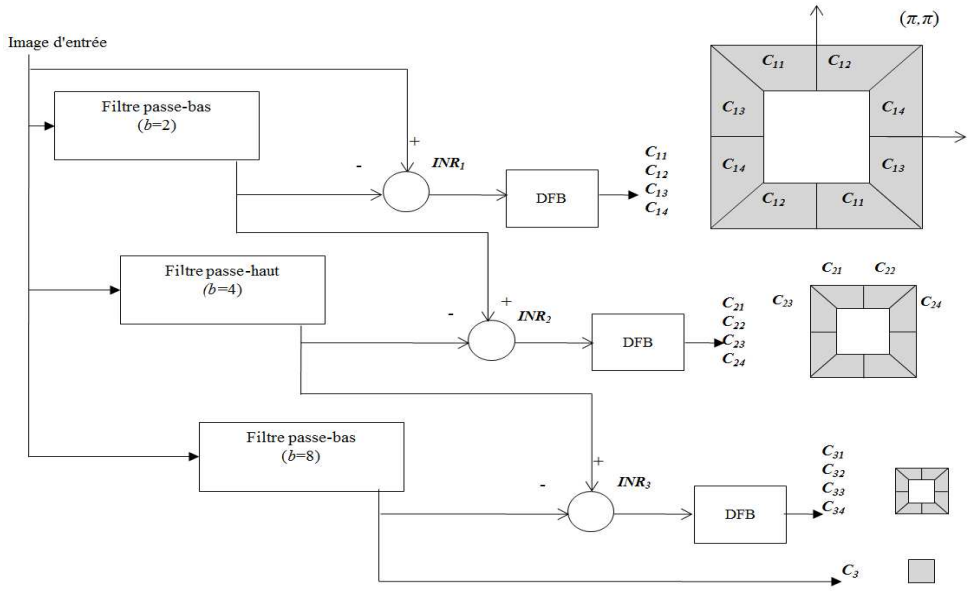


Figure 3.14: Whole diagram of three level RCT decomposition scheme ($L = 3$) and corresponding frequency partition [108].

Fig. 3.15 presents an example of image decomposition that compares between the CT and RCT. We can clearly notice that the RCT has a richer description than the CT in the different subbands. For example, contrarily to the CT, the size of subband images in lower (coarse) levels is the same as for the first level in the RCT, which is very important propriety for texture analysis.

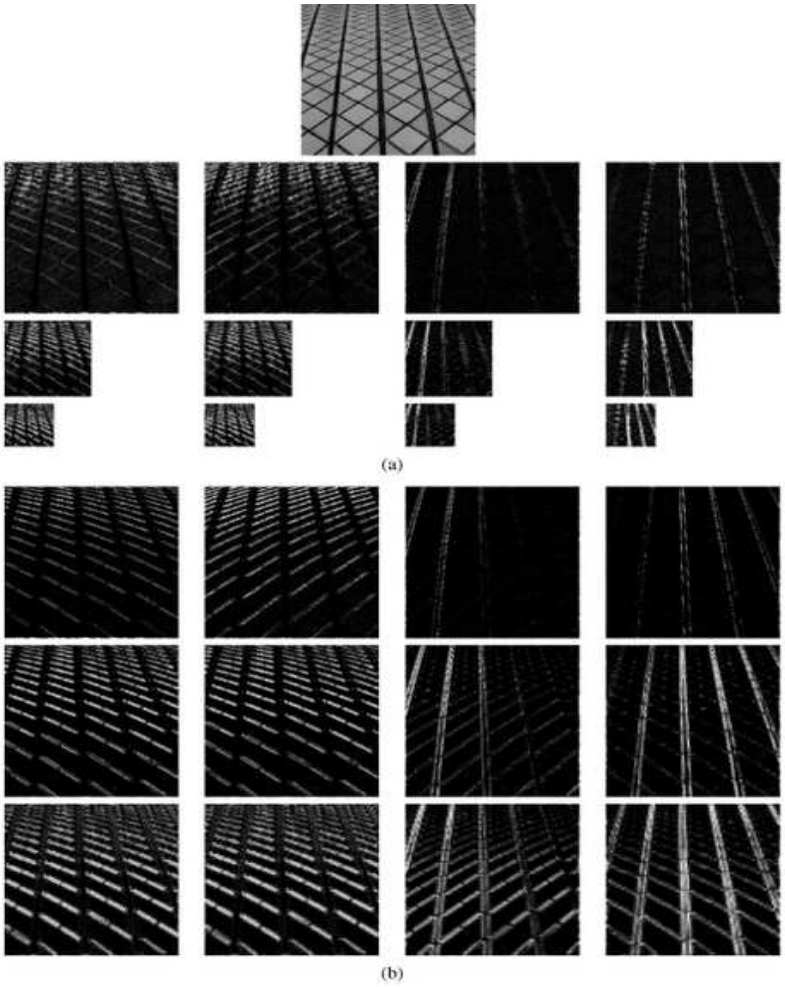


Figure 3.15: Example comparing the CT (a) with the RCT (b). The first row represents the texture image (ref. Tile 1 in the VisTex database [3]). For each transform, from left to right, we show subbands $C_{l,1}, C_{l,2}, C_{l,3}$ and $C_{l,4}$, $l \in 1, 2, 3$, respectively. From top to bottom, we show subband levels $C_{1,d}, C_{2,d}$ and $C_{3,d}$, $d \in 1, 2, 3, 4$, respectively. [109]

3.5 Conclusion

We have presented in this chapter different types of multiresolution transforms used for texture representation: wavelets, contourlets and redundant contourlets. Although wavelets are still a powerful tool in the processing and texture analysis, they admit limits to capture certain directionalities. The contourlets has improved the wavelets capability in this aspect. Therefore, the contourlets provide a good support for multiresolution analysis, thanks to their flexibility to choose the directionality of the subbands. Thus, the detection of smooth contours has been improved.We have also presented the redundant contourlet transform, a variant of the standard contourlets transformation which allow to have subbands of the same size. Thus, we obtain more coefficients facilitating the hierarchical processing of the image, leading to the stability of the modeling.

Chapter 4

Statitistical parametric models for texture representation

4.1 Univariate models

Many kinds of statistical models have been applied to texture representation, but most successful models are generative ones, which describe the texture as a realisation of a random process that can be described by a probability law. Recently, several works have proposed generative modeling of wavelet and contourlet distributions for texture characterization [23, 17, 26]. Since the shape of wavelet coefficients histogram is a representativecharacteristic for the image content, it is important to have flexible statistical models to capture the various histogram shapes [23]. Hereafter, some of these distributions are presented.

4.1.1 Generalized Gaussian density (GG)

The most popular parametric model investigated so far is the univariate generalized Gaussian (GG) density, which has been successfully used for texture classification and retrieval in [23, 24]. Compared with the Laplacian or Gaussian distributions, the GG density has

an additional free parameter that controls the shape of the distribution, which provides flexibility to fit platykurtic and leptokurtic shapes of data [110, 111]. Due to this property, the GG has been also successfully used in image denoising [112], image and video coding [113, 114, 115], image segmentation [116, 89, 7], and texture discrimination and retrieval [23, 117]. Do and Vetterli [23] demonstrated the superiority of using the GG signature over energy-based methods for texture discrimination and retrieval. Other useful applications of the GG also include, speech recognition [118], blind source separation (BSS), [119, 120]. Below the univariate GG formalism.

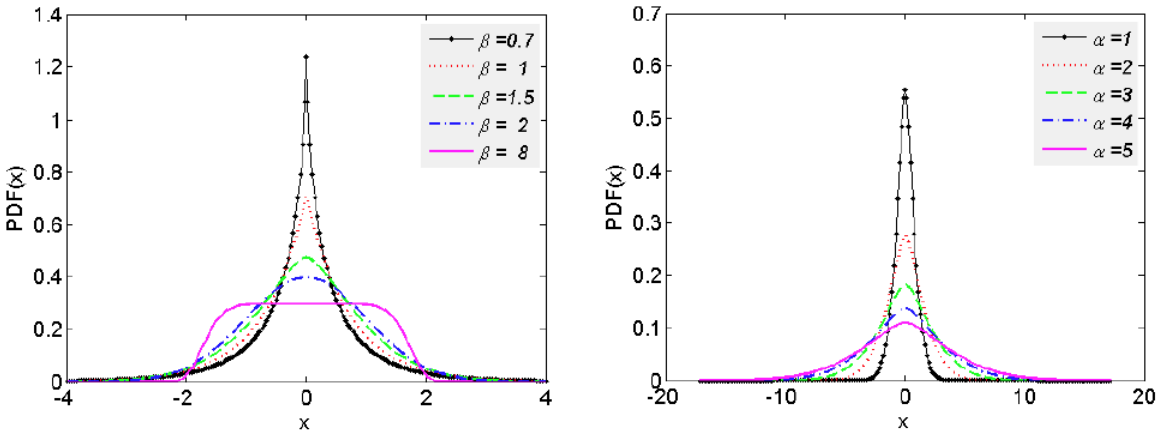
Let x be a univariate random variable with values in $\mathbb{R}$. If x follows a generalized Gaussian distribution, its PDF in its general form [111] is given by:

$$p(x|\mu,\alpha,\beta) = \frac{\beta\sqrt{\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}}}{2\alpha\Gamma\left(1/\beta\right)}\exp\left(-A(\beta)\left|\frac{x-\mu}{\alpha}\right|^{\beta}\right), \tag{4.1}$$

where $\alpha > 0$, $\beta > 0$. $A(\beta) = \left[\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}\right]^{\frac{\beta}{2}}$, $\Gamma(\cdot)$ denotes the gamma function. μ and α are respectively the mean and the standard deviation of the distribution. α is the scale parameter which is related to the width of the peak of the distribution. β is the shape parameter that allows fitting the distribution flattening and thus models its shape in a flexible way. What makes the GGD appropriate in so many applications is its flexible form parameter β which adapts to a large family of symmetric distributions, from super-Gaussian to sub-Gaussian. When $\beta \to \infty$ the distribution becomes uniform on $\left[-\alpha\sqrt{\frac{3\Gamma(3/\beta)}{\Gamma(1/\beta)}}, \alpha\sqrt{\frac{3\Gamma(3/\beta)}{\Gamma(1/\beta)}}\right]$, whereas in the case where $\beta \to 0$, the distribution becomes a delta function centered around μ. Well-known special cases of the General Gaussian Density (GG) function include a Laplacian distribution $\beta = 1$ and Gaussian $\beta = 2$. In the particular case where X follows a zero mean GG law, its PDF is given by,

$$p(x|\sigma,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} exp\left(-\left(\frac{|x|}{\alpha}\right)^{\beta}\right) \qquad x \in \mathbb{R} \tag{4.2}$$

when $\beta = 2$ and $\beta = 1$ zero mean GGD distribution becomes are the Gaussian and Laplacian distributions respectively. Fig. 4.1 (a) shows the various forms of GG of unit variance for different values of β. Fig. 4.1 (b) illustrates the effect of the scale parameter α on dispersion around zero for a GG with a given β shape parameter. In effect, smaller values of the shape parameter β correspond to heavier tails and therefore to more peaked distributions. There are three parameters needed to be estimate in 4.1 which includes the most important parameter β. μ and α can be obtained by the sample average and sample variance respectively, while for the most important parameter (β)the exist many method to estimate it. A review of β estimation can be seen in [121].



(a) (b)

Figure 4.1: (a) The probability density of the generalized Gaussian plotted for different values of the shape parameter $\beta \in 0.7,\ 1,\ 1.5,\ 2,\ 8$. All distributions are normalized to unit variance. (b) The effect of the scale parameter α for GGs with a shape parameter β set to 1.5.

4.1.2 Generalized Gamma density

Recently, texture representation has been improved by replacing the GGD by the generalized Gamma distribution model (GΓD) for wavelet subband characterization [87, 122]. GΓD (see Fig. 4.2) can only model positive values of the coefficients, so absolute value of coefficients is needed to achieved modelisation with this method, consequently one can observe a certain imprecision. In addition to the GΓD used is unimodal.

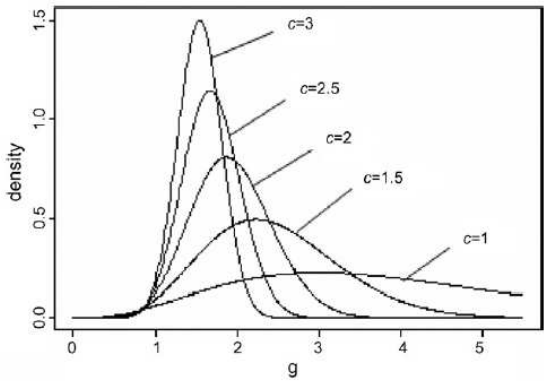


Figure 4.2: Density function for a generalized gamma distribution whith parameter c = 1, 1.5, 2, 2.5, 3 from right to left.

4.2 Mixture of generalized Gaussians (MoGG)

Finite mixture densities allow modelling naturally observations that are supposed to have been generated by a set of different random sources. This allows to perfectly captured the exact forms of the data distributions. By combining the properties of the individual probability density functions, mixture models are capable of approximating any arbitrary distribution. Consequently, finite mixture models are a powerful and flexible tool for modeling complex data. In case of multimodal or suspected multimodal data, the marginal distribution of a random variable $x \in \mathbb{R}$ which is assuming or suspected to be a mixture of

densities (MoGGDs in our case) with components, K is given by:

$$p(x|\vec{\theta}) = \sum_{i=1}^{K} \pi_i p(x|\mu_i, \sigma_i, \beta_i), \tag{4.3}$$

where $0 < \pi_i \leq 1$ and $\sum_{i=1}^{K} \pi_i = 1$. $\vec{\theta}$ denotes the set of model parameters $\{\pi_i, \mu_i, \sigma_i, \beta_i, i = 1, \ldots, K\}$. The goal in modeling is to find the parameters of a model that best fits the shape of the data while keeping a reasonable complexity of the latter (number of components). The parameters are as follows:

- K : Number of components. That is, the number of generalized Gaussians in a model.
- π_i : weight of the component i.
- μ_i : mean of the component i.
- σ_i : standard deviation of the component i.
- β_i : shape parameter of the component i.

For estimate the optimal number of components K, the modeling algorithm uses Minimum message length (MML) [123], [124]. The MML principle provides an optimal way for model selection, providing a good compromise between model complexity and data fit [89, 124]. MML is an optimal way for coding data, because it selects the model with its parameters that gives the shortest overall message length. A more complex model needs a longer message length to encode model parameters, but gives a better fit for the data. On the other hand, a simple model needs a shorter length, but may decrease the accuracy of data fitting. Therefore, the MML provides a natural tradeoff between model complexity and goodness of fit [7]. Given a data sample $\mathcal{X} = \{x_1, x_2, ..., x_n\}$. To estimate the optimal MoGG model, which provides the best tradeoff between data fitting accuracy and model complexity (i.e., the number of components). The message length that encodes the wavelet coefficients in a given subband is given by (see [43] and [54])

The MML of a set of data $\mathcal{X} = \{x_1, x_2, ..., x_n\}$ is given by:

$$MessL \simeq -\log p(\vec{\theta}) + \frac{1}{2}\log|I(\vec{\theta})| + \frac{c}{2}(1 + \log\frac{1}{12}) - \log p(\mathcal{X}|\vec{\theta}), \tag{4.4}$$

where, $p(\vec{\theta})$, $I(\vec{\theta})$ and $p(\mathcal{X}|\vec{\theta})$, denote respectively, the prior distribution of the parameters $\vec{\theta}$, the (expected) fisher information matrix and the data likelihood. $|I(\vec{\theta})|$ denotes the determinant of the matrix $I(\vec{\theta})$. The constant $c = 4K$ the total number of model parameters. Number 4 is the number of parameters (of the set $\vec{\theta}$) per component of generalized Gaussian density. $\vec{\theta}$ parameters initialization is done as indicated in [7].

4.3 Multivariate models

4.3.1 Probabilistic principal component analysis

Probabilistic principal component analysis (PPCA) [125] is a statistical method that estimates the latent variables which generate an input data. PPCA is similar to deterministic principal component analysis (PCA) [126] and often provides the same result. However, unlike deterministic PCA, PPCA is a probabilistic model or generative model which assumes that the input data would be generated under the Gaussian distribution and the Gaussian noise. The extension of plain PCA to PPCA conveys additional practical advantages [125] as follows.

(a) Multiple PCA models may usefully be combined as a probabilistic mixture and how PCA projections may be obtained when some data values are missing.

(b) For dimensionality reduction, probabilistic PCA can be utilized as a constrained Gaussian density model. The benefit of so doing is that the maximum likelihood estimates for the parameters associated with the covariance matrix can be efficiently computed from the data principal components.

4.3.2 Hierarchical probabilistic principal component

Hierarchical probabilistic principal component analysis (HPPCA) [88] is an extention of PPCA (Fig. 4.3 shows a schematic diagram of the HPPCA).

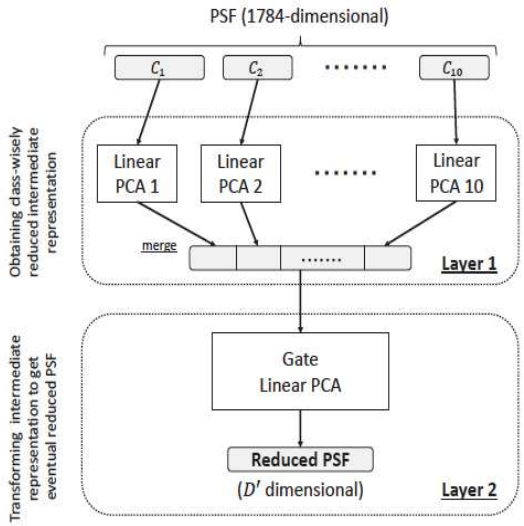


Figure 4.3: Schematic diagram of the hierarchical probabilistic principal component analysis. Input will be divided into known groups and contracted with distinct linear-PPCA. After that, the final linear-PPCA makes a conclusive reduced representation

The HPPCA applies a hierarchical dimension reduction which is based on a known structure or correlation of the input data namely Portilla-Simoncelli statistics (PSS) [22]. Suzuki et al [88] have shown that HPPCA is easier to grasp the latent structure of the input space because of the plain PPCA could not grasp the latent variables of the PSS sufficiently and so, using known group structure of the Portilla's features makes dimension reduction more efficient. The schematic diagram shows that, first, the HPPCA applies a dimension reduction for each classes $C1 \sim C10$ with distinct PPCA models. Second, reduced representations will be concatenated into an intermediate vector, finally, conclusive reduced representation

4.3.3 Multivariate generalized Gaussian density (MGG)

Despite their effectiveness for classification and retrieval of grey scale texture, univariate modeling do not fully describe the statistical behavior of the coefficients because they do not include information on the spatial dependences between the coefficients. To exploit this spatial information, multivariate models are needed to describe the joint distributions of transforms coefficients. *Multivariated generalized Gaussian distributions* (MGG) has been proposed recently in the image processing field, specifically in the analysis of multiresolution transforms coefficients [27]. For instance, it is well known that there is a very important intraband spatial dependence in color image wavelet subbands which can not be captured by univariate models.

To take advantage of the intrinsic information of this dependence, multivariate models should be used to describe the joint statistics of subband coefficients of transforms. The passage from the univariate version to the multivariate one for the joint fitted of the multicomponents is not trivial. Paradoxically, there does not seem to be a general agreement on the multivariate extension of the univariate generalized Gaussian distribution. However we define the centred multivariate GGD (MGG) by the following density:

$$f(\mathbf{X}|\mathbf{\Sigma},\beta) = \frac{\Gamma(\frac{p}{2})\beta}{\pi^{\frac{p}{2}}\Gamma(\frac{p}{2\beta}) \times 2^{\frac{p}{2\beta}}|\Sigma|^{\frac{1}{2}}} \times exp\left\{-\frac{1}{2}\left[\boldsymbol{x}'\mathbf{\Sigma^{-1}}\boldsymbol{x}\right]^{\beta}\right\} \tag{4.5}$$

where p is the dimensionality of the probability space (e.g. $p = 3$ for the three color space bands of a textured color image). $\mathbf{\Sigma}$ a positive define scale matrix. The scalar β is the shape parameter, which controls the peakedness of the distribution and the heaviness of its tails. When $\beta = 1$, we have the multivariate Gaussian distribution, when $\beta = 1/2$, we have the multivariate Laplacian distribution.

Due to the complexity of developing probabilistic metrics in the multivariate case, there

is little work that has involved these models. Recently, considerable attention has been paid to use the MGGD for transform coefficients modeling. Verdoolaege et all ([25], [26], [27]) and [127] have used MGGD for texture retrieval. However, a key limitation of these works is that these multivariate models were mainly applied on wavelet decomposition subbands and their performance evaluated in almost every work in indexing and content-based image retrieval contexts with sometimes too few samples from textures database [34] and [3]. Another limitation concerning these works is that they do not adressed multimodality in texture.

4.3.4 Parameter estimation of the MGG

Note that Gaussian and Laplacian distributions as special cases of MGG. MGG are potentially interesting for modeling the statistical properties of various images or features extracted from these images. In particular, the distribution of wavelet or contourlet coefficients has been shown to be modeled.

Considering the important attention devoted to MGGs, estimating the parameters of these distributions is clearly an important issue. Classical estimation methods that have been investigated for univariate GGDs include the maximum likelihood (ML) method [128] and the method of moments [129]. In the multivariate context, MGG parameters can be estimated by a least-squares method as in [130] or by minimizing a χ^2 distance between the histogram of the observed data and the theoretical probabilities associated with the MGG [131]. Estimators based on the method of moments and on the ML method have also been proposed in [27], [132].

Several works have analyzed covariance matrix estimators defined under different modelling assumptions. On the one hand, fixed point (FP) algorithms have been derived and analyzed in [133], [134] for spherical invariant random vectors (SIRVs). On the other hand,

in the context of robust estimation, the properties of M-estimators have been studied [135]. M-estimators are a broad class of extremum estimators for which the objective function is a sample average. Both non-linear least squares and maximum likelihood estimation are special cases of M-estimators. The definition of M-estimators was motivated by robust statistics, which contributed new types of M-estimators. The statistical procedure of evaluating an M-estimator on a data set is called M-estimation. Based on [135], Pascal [127], shows that the MLE of MGGD parameters exists, and it is unique and can be computed it by an FP algorithm. This latter author establish successfully some properties related to the FP equation of the MLE for MGGs.

All of these estimation methods estimate a single shape parameter for all dimensions which is a limitation. The improvement that we will bring in our contribution will be to have for each dimension a shape parameter in order to realize a more precise estimate.

4.3.5 Summary and discussion

Despite of the univariate models perform well in the areas where they were applied, they share a same limitation, namely that it fails to accurately estimate distributions of the transform coefficients of natural images that can be multimodal [89]. Indeed, it can be observed that the frequency subband distribution for a wide range of natural images is symmetrical and sharply peaked around zero [15, 23, 87]. One can also find a range of images where this distribution is asymmetrical and/or multimodal, as is the case for images exhibiting approximate periodicities [24, 136]. Cossu et al. [136] have observed that wavelet packets histograms can have Gaussian, leptokurtic, or multimodal forms. Using a single GGD or any unimodal distribution to model the wavelet coefficients, in this case, would not perfectly capture the shape of the coefficients distribution, which, in turn, may deteriorate the performance of the wavelet signature for texture discrimination and retrieval. Hence, Since the shape of the wavelet histogram is a critical characteristic for the image content, it is important to have flexible statistical models to capture the various

histogram shapes.

The MoGG models come to some extent to solve this problem capturing multiimodality. However MoGG is designd for univariate date. So, the common limit of these works that used the MoGG models for multicomponent textures representation is that they failed to adress multicomponent and multivalued color texture. So they failed to address simultaneously multicomponent and multidimensionality in texture representation and modelling. To represent and model more accurately the color texture, it is necessary to use a suitable full covariance. That'is another objective of our work where, we propose a new statistical framework based on finite mixtures of multivariate generalized Gaussian distribution (MoMGG).

4.4 Similarity measurement between parametric distributions

There are many similarity measures that are used to discriminate between two distributions. Among them there is the *Kullback-leiber divergence* (KLD). To measure the difference between two probability distributions over the same variable x, the KLD similarity measure, has been popularly used in literature. The concept was originated in probability theory and information theory. [137]

The KLD, which is closely related to relative entropy, information divergence, and information for discrimination, is a non-symmetric measure of the difference between two probability distributions $p_1(x)$ and $p_2(x)$. Specifically, the Kullback-Leibler (KL) divergence of $p_2(x)$ from $p_1(x)$, denoted $KL(p_1, p_2)$, is a measure of the information lost when $p_2(x)$ is used to approximate $p_1(x)$.

Let $p_1(x)$ and $p_2(x)$ are two probability distributions of a discrete random variable $\mathbf{X}$. That is, both $p_1(x)$ and $p_2(x)$ sum up to 1, and $p_1(x) \geq 0$ and $p_2(x) \geq 0$ for any x in $\mathbf{X}$, then

$KL(p_1, p_2)$ is defined by

$$KL(p_1||p_2) = \sum_X p_1(x|\theta_1) ln \frac{p_1(x|\theta_1)}{p_2(x|\theta_2)} \tag{4.6}$$

The KLD measures the expected number of extra bits required to code samples from $p_1(x)$ when using a code based on $p_2(x)$, rather than using a code based on $p_1(x)$. Typically $p_1(x)$ represents the "true" distribution of data, observations, or a precisely calculated theoretical distribution. The measure $p_2(x)$ typically represents a theory, model, description, or approximation of $p_1(x)$. The continuous version of the KLD is given by:

$$KL(p_1||p_2) = \int p_1(x|\theta_1) ln \frac{p_1(x|\theta_1)}{p_2(x|\theta_2)} dx \tag{4.7}$$

Although the KLD measures the "distance" between two distributions, it is not a distance measure. This is because that the KLD is not a metric measure. It is not symmetric: the KLD from $p_1(x)$ to $p_2(x)$ is generally not the same as the KLD from $p_2(x)$ to $p_1(x)$. Furthermore, it need not satisfy triangular inequality. Nevertheless, $KL(p_1||p_2)$ is a non-negative measure. $KL(p_1||p_2) \geq 0$ and $KLD(p_1||p_2) = 0$ if and only if $P_1 = P_2$.

4.4.1 KLD for GGs

Do *et al.* [23] have used a closed-form Kullback-Leibler divergence (KLD) to measure similarity between two univariate GGDs. Closed-form KLD distance between two centred univariate GGDs characterized by β_1, σ_1 and β_2, σ_2 is given by [23] :

$$KL(\beta_1, \sigma_1||\beta_2, \sigma_2) = ln \left[\frac{\beta_1 2^{\frac{1}{2\beta_2}} \sigma_2 \Gamma(\frac{1}{2\beta_2})}{\beta_2 2^{\frac{1}{2\beta_1}} \sigma_1 \Gamma(\frac{1}{2\beta_1})} \right] + \left(\frac{2^{\frac{1}{2\beta_1}} \sigma_1}{2^{\frac{1}{2\beta_2}} \sigma_2} \right)^{2\beta_2} \frac{\Gamma(\frac{2\beta_2+1}{2\beta_1})}{\Gamma(\frac{1}{2\beta_1})} - \frac{1}{2\beta_1} \tag{4.8}$$

where β_1, β_2 the shape parameters, and σ_1, σ_2 the scale parameters.

4.4.2 KLD for MoGGs

When these distributions are multimodal, a closed-form solution is intractable [138]. To circumvent this issue, we resort to approximating the KLD using Monte-Carlo sampling methods, as proposed by [109]. Given two MoGG models $p_1(x) = \sum_{i=1}^{K} \pi_i p_1(x|\theta_i)$ and $p_2(x) = \sum_{j=1}^{M} \omega_j p_2(x|\alpha_j)$, the KLD between these models is defined as follows:

$$KL(p_1||p_2) = \int p_1(x) \log \left(\frac{p_1(x)}{p_2(x)} \right) dx. \tag{4.9}$$

Since there is not closed form for eq.(4.9), we resort to Monte-Carlo integration is given by:

$$KL_{mc}(p_1||p_2) = \frac{1}{n} \sum_{i=1}^{n} \log \frac{p_1(x_i)}{p_2(x_i)} \approx_{n \to \infty} KL(p_1||p_2) \tag{4.10}$$

The approximation method by sampling aims to generate a sufficiently large sample $\mathcal{X} = \{x_1, x_2, ..., x_n\}$ drawn independently from P in order to approximate the KLD using the Monte-Carlo integration [139].

4.4.3 KLD for MGGs

A KLD closed-form between two centered Multivariare generalized Gaussian distributions p-variate (case where $\beta = 1$) was developed by [137].

$$KLD(\Sigma_1||\Sigma_2) = \frac{1}{2} \left[ln \frac{|\Sigma_2|}{|\Sigma_1|} + tr(\Sigma_2^{-1}\Sigma_1 - p \right] \tag{4.11}$$

with Σ_1 and Σ_2 are respective covariances. The closed-form extension to the general case between two MGGs of zero means (where $\beta_1 \neq \beta_2 \neq 1$)) is not easy, however [25] have

been able to obtain a closed-form for this KLDs between two bivariate GGDs.

$$KLD(\beta_1, \Sigma_1 || \beta_2, \Sigma_2) = ln \left[\frac{\Gamma(\frac{1}{\beta_2})}{\Gamma(\frac{1}{\beta_1})} 2^{(\frac{1}{\beta_2} - \frac{1}{\beta_1})} \left(\frac{|\Sigma_2|}{|\Sigma_1|} \right)^{\frac{1}{2}} \frac{\beta_1}{\beta_2} \right]$$
$$+ \left[-\frac{2^{\frac{1}{\beta_1}}}{\beta_1} + 2^{\frac{\beta_1}{\beta_2}} \frac{\Gamma(\frac{\beta_2+1}{\beta_1})}{\Gamma(\frac{1}{\beta_1})} \times \left(\frac{\gamma_1 + \gamma_2}{2} \right)^{\beta_2} 2^{\mathbf{F}\left(\frac{1-\beta_2}{2}, -\frac{\beta_2}{2}; 1; A^2\right)} \right] \tag{4.12}$$

where $\gamma_i \equiv (\lambda_2^i)^{-1}$, $i = 1, 2$, λ_2^i the eigenvalues of $\Sigma_1^{-1}\Sigma_2$. $A \equiv \frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2} \cdot 2^{\mathbf{F}(.,.;.;.)}$ the hypergeometric function of Gauss which can be tabulated for realistic values of β and $-1 < A < 1$. For $\beta_1 = \beta_2 = 1$ the hypergeometric function becomes 1 and 6.17 is reduced to 6.18. However, It was impossible to find closed-form for for the KLD between MGGs of zero means for a dimension over to 2.

4.4.4 MoMGG expected potential

The Table 4.1 exhibits a comparison of well-known methods of texture representation in the literature with the potential of our expected unified MoMGG model; subsequently in the pursuit of our contribution, we will demonstrate theoretically and practically its effectiveness. The comparison criteria being the types of data processed. Clearly, we can observe that our method is the only one which could take into account any type of distribution, whether univariate, unimodal, multivariate and/or multimodal.

	Univariate	Multivariability reliying on joint marginal distributions	Multivariability reliying on the full covariance	Multimodality
MoMGG	**X**	**X**	**X**	**X**
MoGG	X	X		
MGG	X		X	
GG	X			
SIRV	X	X		
Copulas	X	X		

Table 4.1: Comparison of capabilities to address different types of data distributions

4.5 Conclusion

In this chapter we have reviewed some relevent statistical parametric models of texture representation. We represented the most popular models investigated so far, namely the univariate and MoGG models commonly used to statistically describe data. For univariate data modeling case, we have shown for multimodal, asymmetric, non-Gaussian data, that a single GG cannot correctly fit these data, and that mixture models are the most appropriate. To represent multimodal data, univariate MoGG has received a lot of attention in the last decade. It have been successfully applied to monochrome texture discrimination and retrieval, fabric texture defect detection and IR face recognition. Multivariate (MGG) modelling has also received much attention in the literature to model multivariate texture data, particularly in the image processing community even if its parameters estimation require many restrictive assumptions.

Multivariate finite mixture of generalized multivariate Gaussian modeling (MoMGG) has never been investigate. So, there is clearly a lack of a unified statistical framework integrating multimodal and multidimensional data. Therefore, an efficient formalism of MoMGG parameters estimations is still lacking. We also reviewed a particular similarity

mesure between GGDs and MGGs. Concerning the distance between MGGs, we note that, not only, it has not been possible in the previous work to determine an analytical nor an approximate distance between more than two MGGs. Also, there is no work dealing with distance measurements between MoMGGs.

Chapter 5

Automatic fabric Defect Detection Using Learning-Based Local Textural Distributions in the Contourlet Domain

5.1 Motivation

In this first contribution, we will apply the MoGG in the context of the binary classification of textures. More specifically, we use MoGG to detect defects in texture images. We applied the siganture to image defect detection in local block level in a Bayesian learning-based approach. Our approach is based on a local statistical representation of fabric patterns using the redundant contourlet transform (RCT). The distribution of the RCT coefficients are modeled using finite mixtures of generalized Gaussians (MoGG), which constitute statistical signatures distinguishing between defected and defect-free fabrics. In addition of being compact and fast to compute, these signatures enable accurate localization of defects.

Our defect detection system is based on three main steps. In the first step, a preprocessing is operated on the fabric image to detect basic pattern size for image decomposition and signature calculation. In the second step, labeled fabric samples are used to train a Bayes classifier (BC) to discriminate between defect-free and defected fabrics. Finally, defects are detected on inspected new images by testing local patches using the learned BC. Our approach can deal with multiple types of textile fabrics, from simple to more complex ones. Experiments on the TILDA databaset [140] have demonstrated that our method yields better results compared to recent state-of-the-art methods.

5.2 Introduction

Textile is used in multiple products such as clothing, filters, wipes, and in housing and transportation materials. However, the presence of defects in fabrics can reduce prices with losses reaching 45%-65% [142]. To enhance the efficiency of fabric defect detection, it is necessary to replace the fastidious manual inspection with automatic inspection for better productivity and improving quality of fabric as well (see Fig. (5.1) as an example for automatic inspection systems).

Automated defect detection of fabrics reduces the labor cost and enables to cover a broader range of different fabrics, from homogenous texture to the most complex one. Currently, there exist more than 70 established categories for fabric defects defined by the textile industry [143]. Most of these defects are caused by machine malfunctions, yarn problem, stain of oil caused by the knitting device, among others.

Fabric can be considered as a two-dimensional (2D) patterned texture [143, 144]. All fabrics can be classified into the 17 established wallpaper groups denoted as $p1$, $p2$, $p3$, $p3m1$, $p31m$, $p4$, $p4m$, $p4g$, pm, pg, pmg, pgg, $p6$, $p6m$, cm, cmm and pmm, that are lattices composed of elementary elements called *motifs* which are organized repetitively along par-

Figure 5.1: Example of an automated inspection machine for textile fabric.

allelogram, rectangular, rhombic, square or hexagonal shapes [145]. The $p1$ group defines a texture with just one fundamental lattice repeating itself over the complete fabric such as plain, twill and plain weave fabrics (see Fig. (5.2) for illustration). This group involves only pattern translations. The other groups involve one or more other types of symmetries such as rotations, reflections and glide-reflections (see Fig. (5.3) for illustration). Considering the classification of inspected fabric types, Ngan et *al.* [143] have proposed a taxonomy for defect detection methods which broadly categorizes the methods into two main groups. The first group of methods (non-motif-based), do not take into consideration the basic motif distribution for fabric inspection. They rely instead on analysing textural properties of fabric images. These methods can, therefore, deal with fabrics of type $p1$ (e.g., plain and twill) as well as with other types of fabrics in the wallpaper groups. The second group of methods (motif-based) relies on the analysis of the structure of the fabric basic motifs to detect potential defects. Since it is very difficult to model statistically/geometrically the distribution of basic motifs in these groups, methods use mainly pattern matching and

image subtraction to localise potential defects [146].

Several fabric defect detection methods have been proposed in the last three decades [130, 142, 143]. The majority of these methods is specific to the non-motif-based category and deal mainly with fabrics of type $p1$ [147, 148, 149, 150, 151, 152]. Recently, some approaches have been proposed for motif-based defect detection [150, 146]. These have been evaluated mainly on texture patterns of $p1$ and pmm fabric types. However, they do not generalize well for other types of fabric. In addition, these methods detect defects at the image level (i.e., an image contains a defect or not) and do not provide defect details at the lattice or motif level. To resolve the above limitations, some authors have proposed generic methods that can be applied to multiple types of fabrics to localise the defects [147, 153]. These methods are often based on template-matching or statistical techniques that compare inspected fabric to defect-free reference one to detect defects [148, 143]. Ngan *et al.* [153] have proposed to analyse the energy of motif subtraction to detect defects. This method can be applied to 16 of the 17 wallpaper groups and has shown some robustness to noise and slight motif shift. However, it requires a defect-free ground truth to compare the motifs. In addition, it uses a circular shift to perform image subtraction which incurs a huge computation time. Ng *et al.* [148] have used energy minimization to decompose an image into two structures: a cartoon image representing the defect and the texture structure representing the repetitive patterns. This method yields very good results for defect localisation and can be applied practically to all fabric types. However, it requires the inspected image to be perfectly aligned with the ground truth. In addition, it is not efficient for detecting defects with small contrast in the image (e.g. oil stains, etc.).

In this chapter, we introduce a novel defect detection algorithm which has the capability to cope with different types of defects in the $p1$ and non-$p1$ groups. Therefore, our method can be considered as a hybrid one since it relies mainly on analysing texture patterns for fabric defect detection. Strictly speaking, our method does not consider elementary motifs as a basic manipulation unit [143]. Instead, it decomposes the image into elementary repet-

itive units (ERUs) characterized by their periodicity along the fabric lattice. Contrarily to analysing the fabric at the image level [148] or using the basic motifs [143], ERUs constitute an intermediate representation coarsely describing the local structure of the texture to allow an analysis similar to methods of type $p1$ group to be performed. Each ERU undergoes first a multi-scale Contourlet decomposition describing local directional and structural properties of the fabric texture. A statistical signature of generalized Gaussian distributions (MoGG) is then calculated on the Contourlet coefficients [138, 154], which constitute our representation for fabric comparison. Given a set of labeled samples of signatures representing defect-free and defective fabric ERUs, we train a Bayes classifier (BC) to separate between defect-free and defective fabrics, which is then applied to inspect new images for defect detection. Experiments conducted on the TILDA database [140] have shown that the proposed algorithm yields better results when compared to other previous works. The main contributions of our work in this chapter can be summarized in the following points:

- A generalized non-motif-based method is proposed for fabric defect detection. Our method not only can deal with the $p1$ fabric group, but also with other fabric groups. It relies on an intermediate representation of fabrics using ERUs which facilitates local inspection of images with complex patterns and enables accurate localisation of defects.

- To describe fabric texture structure, we use multi-scale contourlet decomposition (RCT) and mixtures of generalized Gaussian distributions (MoGG) which allow robustness to noise and non-uniform illumination. It has also invariance to fabric translation and scale changes. These descriptors calculated on the ERUs are called RCT-MoGG signatures.

- Our defect detection method is based on a learning-based approach which separates RCT-MoGG signatures of defect-free and defective fabrics using a set of learning examples with their ground-truth. After decomposing a fabric image into ERUs, a

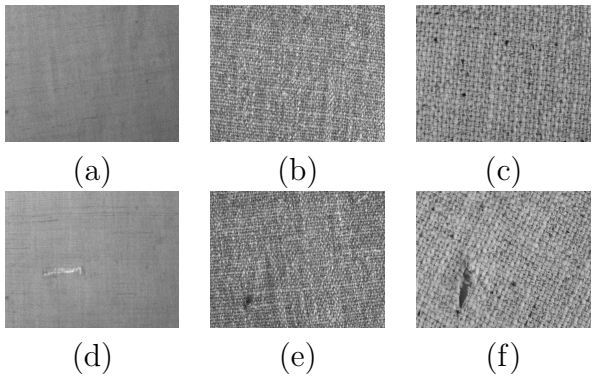


Figure 5.2: Fabric images from TILDA database representing the $p1$ wallpaper group: (a) plain fabric without defects , (b) twill fabric without defects, (c) plain weave fabric without defects, (d) plain fabric with defects, (e) twill fabric with defects and (f) plain weave fabric with defects.

classification of the ERU blocks is performed to assess whether each ERU is defect-free or contains defects. Our method provides an overall high detection accuracy and very low false alarm rate compared to recent state-of-the-art methods.

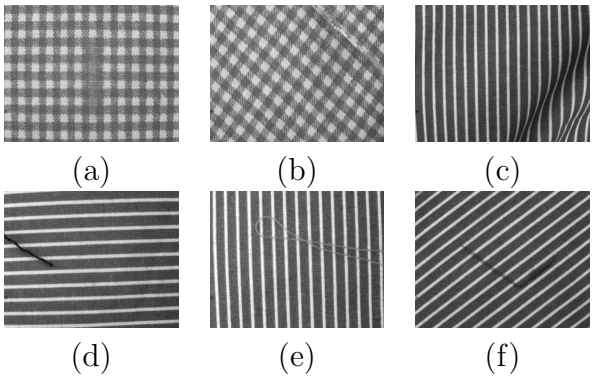


Figure 5.3: Non-$p1$ wallpaper group fabric images in the TILDA database: (a) gingham fabric ($p4m$), (b) gingham oblique fabric, (c) fabric with vertical stripes ($p2$), (d) fabric with horizontal stripes, (e) fabric with vertical stripes and (f) fabric with oblique stripes.

The remainder of this paper is organized as follows: Section 5.3 presents a brief literature review on fabric defect detection methods. Section 5.4 presents our method for automatic defect detection. Section 5.5 presents some experimental results. We end the paper with a conclusion.

5.3 Related work

The majority of existing work on fabric inspection is related to the *non-motif-based* category of methods [142]. These methods have been designed for the simplest patterns of the $p1$ homogenous fabrics [155], which include plain and twill fabrics. Proposed *non-motif-based* methods can be classified into four main approaches: *statistical*, *spectral*, *model-based* and *learning-based* approaches [143]. Compared to the *non-motif-based* category, a small number of methods have been proposed for the *motif-based* one [153].

5.3.1 Inspection approaches for the $p1$ fabrics

Statistical methods include mainly the analysis of *the auto-correlation function* (AF) and the co-occurrence matrix (CM) [156, 186]. The AF has been applied to detect fabric defects by assuming that auto-correlation maxima stay constant for repetitive patterns with perfect replication throughout the fabric [156, 170]. In the same vein, the AF is applied to describe translational and rotational symmetry in plain fabric images [186]. This approach although interesting has not been evaluated by explicit performance metrics on fabric defects [143]. Tsai *et al.* [186] have used features calculated from the CM to classify fabrics as normal or defective using neural networks. Likewise, Amet *et al.* [157] have applied CM analysis in the wavelet domain to detect defects. In [167], CM features are fed into a neural network to detect defects at the pixel level. The major limitation of using CM features is that they incur a huge computation time for their calculation. In addition, since CM computing is based on adjacent pixels analysis, CM features can be sensitive to scale changes in the fabric texture.

Spectral approaches consist of locating defects in the spectral domain [142]. These approaches include the Fourier transform (FT) [158, 183], the wavelet transform (WT) [179, 187, 188] and the Gabor transform (GT) [94, 166, 167]. The main drawback of using the FT is that it lacks local support (i.e., lack of information support in the spatial domain) which can prevent it from detecting defects in random patterned texture such as twill and plain fabrics. Contrarily to the FT, wavelet and Gabor filters make use of a spatial-frequency analysis which enables them to detect local defects. For instance, Chin *et al.* [156] have proposed a wavelet-based approach to detect defects on plain and twill fabrics. The authors reported a detection success rate of 97.5%, but the method is computationally intensive.

Model-based approaches address the defect detection problem by fitting parametric models to feature distribution of the fabric [142, 175]. Inspired by existing works on texture modelling, Cohen *et al.* [161] have used Gaussian Markov fields (GMFs) to model

defect-free texture on fabric images. In the same vein, Ozdemir *et al.* [176] have applied Markov random fields (MRFs) for fabric inspection. Although GMFs and MRFs are good for extracting local contextual information of texture patterns, they are not efficient in detecting small defects. In addition, these methods are computationally intensive. To exploit linear dependency between pixels, Limas *et al.* [173] have used autoregressive models (AR) to detect defects on leather surfaces. Although this method can operate in real-time scenario, it is sensitive to the size of defects and non-uniform illumination.

Learning-based approaches use labeled samples to train classifiers that assign images to defective or defect-free classes. For instance, Kuo *et al.* [169] have used back-propagation (BP) neural networks for defect detection. They achieved very good detection rates on defective samples of plain fabric. This method, however, does not generalize well to other types of fabric and it is computationally intensive. Sezer *et al.* [180] have used independent component analysis (ICA) on a sample of images to detect defects at the block level of fabrics. This approach has yielded good results on uniform textures (plain fabric). However, it does not generalize well for irregular random textures (e.g., twill and plain weave textures).

5.3.2 Inspection approaches for the non-$p1$ fabrics

These are also named *hybrid approaches* since they use a combination of techniques for defect detection on fabrics in the $p1$ and non-$p1$ groups [143]. They can be broadly classified as *template-matching* (TM) and statistical/spectral approaches. Approaches using TM use image subtraction methods to compare inspected fabric to defect-free ones [153]. Kuo *et al.* [147] have proposed a method using features calculated from the co-occurrence matrix (CM) and correlation analysis to detect defects. This method has been successfully applied to Jacquard fabrics belonging to the $p2$ group. However, CM features are highly sensitive to small variation or misalignment in patterns and their calculation is computationally prohibitive. Tajeripour *et al.* [181] have proposed a method based on local binary patterns

(LBP) and adaptive thresholding for defect detection. The LBP technique has been first proposed for texture description and has good properties for rotation invariance and multi-scale analysis. This method has shown good results for $p1$ and several non-$p1$ groups, but it may yield too many false positives for erroneous thresholds. Besides, the size of the windrow is not adaptable since it its always fixed to 16×16.

Direct thresholding (DT) [149] has been used for the analysis of horizontal and vertical sub-bands of the Haar wavelet transform for defect detection. This method named wavelet golden image subtraction (WGIS) is computationally fast and leads to good performance on several types of fabrics. However, the calculated thresholds can be hard to determine for some fabrics like charter-box pattern. Recently, Tsang *et al.* [152] have proposed a fabric inspection method based on the Elo-rating formalism [163]. This method can effectively deal with non-$p1$ fabric groups like dot-patterned (pmm group) and star-patterned ($p2$ group) and box-patterned fabrics ($p4m$ group). However, it showed weak generalization to other types of fabric. Bollinger bands (BB), originally introduced for financial analysis, have been successfully applied for defect detection in fabric groups $p2$, pmm and $p4m$ [150]. But , defects near the image border or smaller than one repetitive unit may not be detected by the method.

Recently, Ng *et al.* [148] have proposed an energy minimization method for decomposing an image into defect and normal fabric images. This method can be applied for several types of fabrics. However, the method can be very sensitive to noise, defect contrast and errors of alignment between reference (defect-free) and inspected fabrics. Finally, sparse coding (SC) is another successful technique that has been used for defect detection [189, 190]. SC aims at approximating an input signal as a linear combination of a few components selected from a dictionary of basic elements called *atoms* [171]. Zhou *et al.* [190] have proposed to use SC for defect-free fabric reconstruction. Thus, reconstruction error acts as an indicator for the presence of defects in an inspected fabric. The major limitation of the method, however, is its computation time and weak robustness to noise.

To reduce the noise effect, Gabor filters can be used before applying the SC technique as proposed in [191]. The authors have reported a significant improvement on plain, twill and plain weave fabrics from the Tilda dataset [140].

5.3.3 Motif-based methods

Among recent studies for defect detection, only few papers consider elementary fabric motifs as a basic manipulation unit [143]. Ngan *et al.* [153, 146] are among recent works falling into this category of methods. They try to handle explicitly fabrics in the non-$p1$ groups and use symmetry properties of motifs to calculate the energy of moving subtraction and its variance among motifs. Decision boundaries are determined by learning the distribution of those values among the defect-free and defective patterns in the energy-variance space [153]. An extensive performance evaluation on defects in 16 out of 17 wallpaper groups has achieved an overall detection success rate of 93.86%. However, the method is computationally intensive because too large database of defect-free and defective samples is needed for training. Besides, it uses a decision boundary (a sort of threshold) which requires to be manually set for each fabric to guarantee good results, thus the method is not easy to generalize to all types of fabric. To isolate defects on fabric images, Jing *et al.* [165] have proposed to subtract a golden template from a filtered fabric image using Gabor filters. The method has shown good results for defect detection and localisation on patterned fabrics. However, it can be sensitive to noise and non-uniform illumination which increases the amount of false positives. In addition, because of the difficulty of calculating the golden image template, the performance decreases for irregular patterned fabrics.

5.4 The proposed approach for defect detection

The proposed method aims at developing a principled methodology ensuring full automation of defect detection while enabling to efficiently cope with several types of fabrics and

defects. As stated before, our approach is composed of three basic steps. In the first step, a pre-processing is used to determine the ERUs for image decomposition. The second step consists of calculating statistical signatures on labeled samples of fabric images. This step involves also training a Bayes classifier to discriminate between defective and non-defective fabrics. In the third and final step, images of inspected fabrics are passed through the signature generator and the Bayes classifier to detect potential defects. In what follows, we give details about each of the above steps separately:

5.4.1 Image decomposition into ERUs

To make our approach generalizable to multiple fabric groups, we use an intermediate representation of the fabric called elementary repetitive unit (ERU). In a nutshell, an ERU can be considered as a super motif which enables dealing with non-$p1$ fabric group using methods of type $p1$ group. We calculate these units by analysing texture patterns of the fabrics using the autocorrelation function (AF) on the fabric image.

Let I be a fabric image of size $n \times m$. We denote by $I_{hs}(i)$ the image resulting from using i horizonal circular shifts for the image I. Likewise, we denote by $I_{vs}(j)$ the result of using j vertical circular shifts for I. The size of a repetitive pattern (i.e., period) is determined by peaks of the AF of the image I that can be calculated using the original I and its shifts in the horizontal and vertical directions , respectively. The following functions can be used to compute the size of the period in the two directions:

$$\begin{cases} h' &= \arg\max_{1 \leq i < n} \left\{ N(I) \circ N(I_{hs}(i)) \right\} \\ w' &= \arg\max_{1 \leq j < m} \left\{ N(I) \circ N(I_{vs}(j)) \right\} \end{cases} \tag{5.1}$$

where $N(\cdot)$ is a function having the role of normalizing and putting an image in a vector form. The operator $\circ$ indicates the correlation function between vectors. Basically, Eq. (5.1) determines the horizonal and vertical dimensions of the repetitive pattern. Finally, to enable robust RCT decomposition and MoGG parameter estimation, we should maintain a

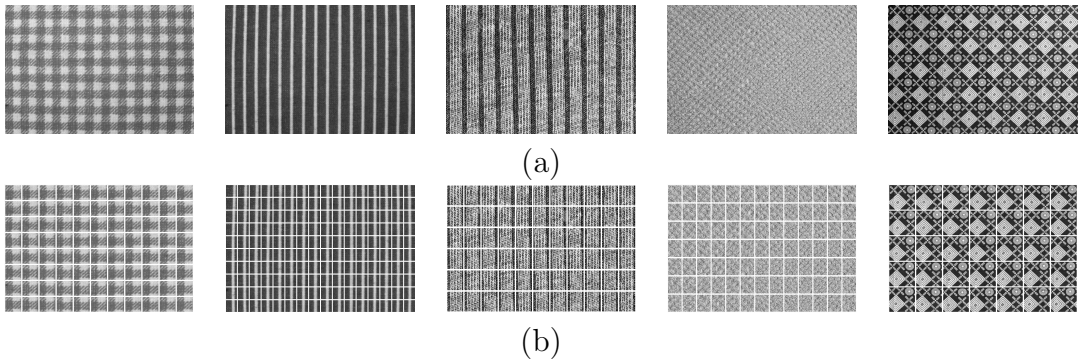


Figure 5.4: Illustration of ERU calculation for some fabric samples: (a) original fabric and (b) image grid with ERUs.

minimum size of the ERU such that a sufficient number of pixels are contained in the unit. By taking δ as the minimum size of each side of an ERU, the optimal height and width of an ERU are given by: $\mathbf{h} = k_h \cdot h'$ and $\mathbf{w} = k_w \cdot w'$, such that $k_h = \{\arg\min_k \left(k \cdot h'\right) | k \cdot h' \geqslant \delta\}$ and $k_w = \{\arg\min_r \left(r \cdot w'\right) | r \cdot w' \geqslant \delta\}$, where $k, r \in \mathbb{N}^*$. Fig. (5.4) gives some examples of ERU calculation on fabric images where the parameter δ is set to 64. We can notice that the periodic patterns have been clearly identified using our method.

Once the size of the ERU is determined, the image is divided into blocks which are used to calculate the statistical signatures for defect detection. In other words, taking the detected ERU as an elementary fabric motif, any fabric can be brought to the $p1$ group, since the only transformation undergone by the ERUs consist of horizontal and vertical translations. Note that having an ERU for each fabric image incurs that before analysing inspected fabric for defect detection, it must be decomposed into ERUs in the same way as the reference fabric. Consequently, in order to have comparable ERU, a preprocessing step might be required to align the inspected image with the reference one.

5.4.2 MoGG modeling for redundant contourlet transform

The standard contourlet transform (SCT) has been proposed [162] to improve representation of texture over wavelets. Later on, the redundant contourlet transform (RCT) has

been proposed as an alternative to the SCT [154] to overcome the limitation due to sub-sampling in the (SCT). In addition of having the good properties of the SCT, the RCT reduces the decimation, avoids the multi-scale interpolation in the SCT and enables easy correspondence between the different levels of image resolutions. The RCT shares the same decomposition scheme with SCT. However, all down-sampling operations in the RCT are discarded from the Laplacian stage and a set of symmetric low-pass filters having adequate frequency selectivity and pseudo-Gaussian properties are employed. Filter impulse responses $g_b(s)$, as given in Eq. (5.2), are finite and symmetric, where s is the spatial row/column location and b is a factor influencing the frequency bandwidth:

$$g_b(s) = e^{-2\frac{s}{b}} - e^{-2}\left(e^{-2(\frac{s-b}{b})^2} e^{-2(\frac{s+b}{b})^2}\right) \tag{5.2}$$

Using L filters (with $b = 2^\ell$, $\ell \in \{1, 2, .., L\}$) results into a redundant Laplacian pyramid (RLP) having $L + 1$ equal-size sub-images: one coarse image approximation and L band-pass sub-images. Then, a directional filter bank (DFB) with $D = 4$ orientations and 1:4 critical down-sampling is applied on each of the L RLP subbands to obtain $4L$ equal-size directional subbands ($C_{ld}, l = 1, \ldots, L;\ d = 1, \ldots, D$) in addition to a 1:4 down-sampled image approximation C_L. The RCT has been successfully used in [159, 109] for texture retrieval and defect detection on plain fabrics.

In the present work, we use the RCT for describing the texture structure of several types of fabrics, from the $p1$ to the non-$p1$ groups. More specifically, let $x_1, ..., x_n$ be the produced RCT coefficients at a given sub-band after decomposition. Since the distribution of these coefficients can be multi-model, we propose to model this distribution using a mixture of generalized Gaussians (MoGG) [7, 160]. By supposing a mixture of K components, the marginal distribution of the coefficient variable $\mathbf{x} \in \mathbb{R}$ is given by:

$$p(x|\Theta) = \sum_{i=1}^{K} \pi_i p(x|\theta_i), \tag{5.3}$$

where $\theta_i = \{\mu_i, \sigma_i, \beta_i\}$ and π_i, $i = 1, ..., K$, are the mixing parameters, with $0 < \pi_i \leq 1$ and $\sum_{i=1}^{K} \pi_i = 1$. Each component of the mixture is modeled using a general Gaussian distribution (GGD): $p(x|\theta_i) = \frac{C(\beta_i)}{2\sigma_i} \exp\left(-A(\beta_i)\left|(x-\mu_i)/\sigma_i\right|^{\beta_i}\right)$, where $A(\beta_i) = \left[\frac{\Gamma(3/\beta_i)}{\Gamma(1/\beta_i)}\right]^{\frac{\beta_i}{2}}$, $C(\beta_i) = \beta_i \sqrt{\frac{\Gamma(3/\beta_i)}{\Gamma(1/\beta_i)}} / \Gamma\left(1/\beta_i\right)$, and $\Gamma(\cdot)$ denotes the Gamma function. The parameters μ_i and σ_i denote the distribution mean and standard deviation, respectively. The parameter β_i is the shape parameter which fits the kurtosis of the ith GGD and determines whether the distribution is peaked or flat. The details of parameter estimation of the MoGG are given in [138]. We use a Bayesian estimation to learn the model parameters θ_i, $i = 1, ..., K$, whereas the minimum message length (MML) principle is applied for selecting the best value of the parameter K [124].

Fig. (5.5) shows comparison of the RCT-MoGG signatures of two blocks in the same fabric (Fig. (5.5).a: a defect-free block (green) and a defective one (red). Note first that the shape of coefficients histogram can be sharply peaked, heavy-tailed and slightly asymmetric. Thanks to its flexibility, the MoGG has enabled to precisely fit to the different shapes. We can note also the discrepancy between the signatures of defective and defect-free blocks for both high- and low-pass subbands. This makes RCT-MoGG description a very efficient tool to assess about defects present in the fabric.

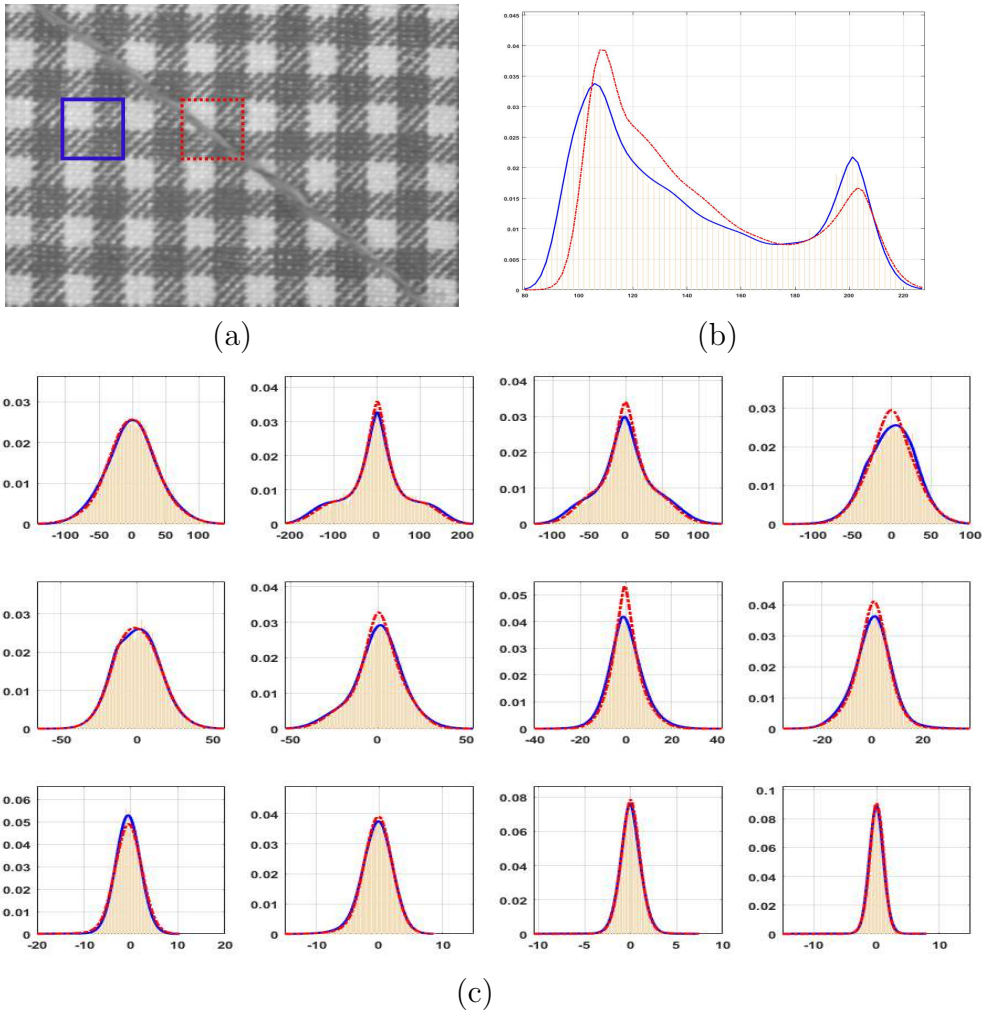


Figure 5.5: RCT-MoGG signature calculation for defective and defect-free fabrics: (a) shows examples of defective (dashed red) and defect-free (solid blue) blocks, In (b) and (c), we show in yellow the histograms of low- and high-pass RCT coefficients of the defect-free block in (a). The dashed-red and solid-blue lines show MoGG signatures of the defective and defect-free blocks, respectively. In (c), from left to right, we show the four directional decompositions, and from top to bottom, we show levels 1, 2 and 3 of the RCT-MoGG signatures, respectively.

5.4.3 Similarity measurement for defect detection

Do *et al.* [162] have used a closed-form Kullback-Leibler divergence (KLD) to measure similarity between two statistical distributions of wavelet subband coefficients. When these distributions are multi-modal, a closed-form solution is intractable [138]. To circumvent this issue, we resort to approximating the KLD using Monte-Carlo sampling methods, as proposed by [109]. Given two MoGG models $P_1(x) = \sum_{i=1}^{K} \pi_i p_1(x|\theta_i)$ and $P_2(x) = \sum_{j=1}^{M} \omega_j p_2(x|\alpha_j)$, the KLD between these models is defined as follows:

$$KLD(P_1||P_2) = \int P_1(x) \log \left(\frac{P_1(x)}{P_2(x)} \right) dx. \tag{5.4}$$

The KLD by Monte-Carlo integration is given by:

$$KLD_{mc}(P_1||P_2) = \frac{1}{n} \sum_{i=1}^{n} \log \frac{P_1(x_i)}{P_2(x_i)} \approx KLD(P_1||P_2). \tag{5.5}$$

The Monte-Carlo method aims at generating a sufficiently large sample $\mathcal{X} = \{x_1, x_2, ..., x_n\}$ drawn independently from the distribution P_1 to approximate the KLD integration [139]. We used $n = 10^4$ as typical value which gives satisfactory results.

5.4.4 Learning-based defect detection algorithm

The flow diagram for our algorithm is shown in Fig. 5.6. As mentioned previously, our approach works into two stages: *a learning phase*, in which a Bayes classifier is trained on a set of labeled fabric examples, and an *inspection phase*, which uses the trained classifier on a set of newly seen images to detect potential defects. Our algorithm operates at the block level of images which aims at localising defects at fine resolutions allowed by the ERU decomposition.

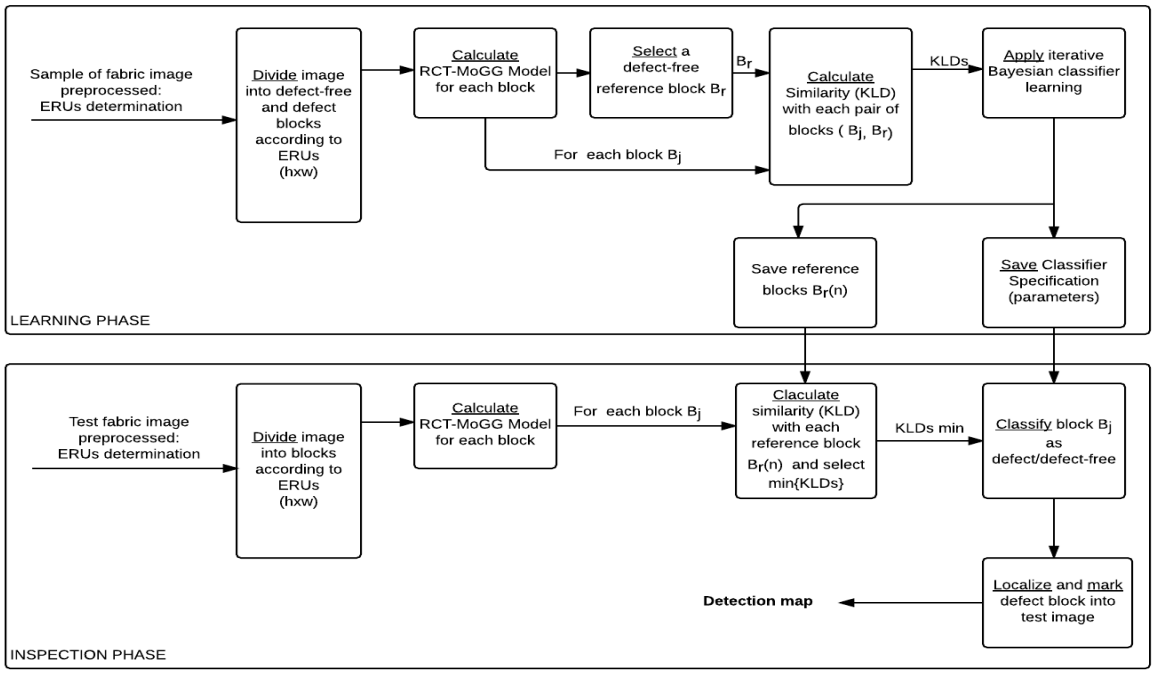


Figure 5.6: Block diagram of the proposed defect detection method.

5.4.4.1 The learning stage

Let $\mathcal{T} = \mathcal{B} \cup \tilde{\mathcal{B}}$ be the set of training examples composed of two subsets $\mathcal{B} = \{B_1, ..., B_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{B}_1, ..., \tilde{B}_m\}$, containing defect-free and defective blocks, respectively. For each block in $\mathcal{T}$, we calculate its RCT-MoGG signature which combines several orientations and scales of the texture. The training process of the defect detection system is given by the script of Algorithm 5.1. The algorithm trains a Bayes classifier BC and results into a set of reference blocks $\mathcal{R}$ that will be used to inspect newly-seen textile images. The set $\mathcal{R}$ is a sort of a landmark signature that contains the different configurations and patterns of the textile type to be inspected.

The training process iterates on the sets $\mathcal{B}$ and $\tilde{\mathcal{B}}$ by choosing at each time a new reference block to add to the set $\mathcal{R}$. The first reference block $B_{r_1} \in \mathcal{R}$ can be chosen randomly from $\mathcal{B}$. After calculating the KLDs of all the training blocks with B_{r_1}, we

obtain the sets $\mathcal{D} = \{d_1, ..., d_n\}$ and $\tilde{\mathcal{D}} = \{\tilde{d}_1, ..., \tilde{d}_m\}$ for $\mathcal{B}$ and $\tilde{\mathcal{B}}$, respectively. Then, we train a Bayes classifier (BC) on $\mathcal{D} \cup \tilde{\mathcal{D}}$ where a Gaussian $P_{\mathcal{B}}$ (resp. $P_{\tilde{\mathcal{B}}}$) is fitted to class $\mathcal{B}$ (resp. $\tilde{\mathcal{B}}$). After classifying the training blocks using BC, we obtain the classification error ε (i.e., number of badly classified blocks) and the set $\mathcal{C}$ of false defect detections (i.e., blocks in $\mathcal{B}$ classified as defect by BC).

To augment the set $\mathcal{R}$ with a new reference block, we search for a block in $\mathcal{C}$ corresponding to either the median or the maximum of distances in $\mathcal{C}$. Using the new set $\mathcal{R}$, we update the distances $\mathcal{D}$ and $\tilde{\mathcal{D}}$ as follows. Let B_k be a block in $\mathcal{T}$ and $KLD_{r_1}, ..., KLD_{r_N}$ are the set of KLDs calculated with all reference blocks in $\mathcal{R}$, where N is the cardinality of $\mathcal{R}$. The new KLD assigned to the block for the next iteration (Step 2 of the algorithm) is given by $d_k = \min\{KLD_{r_1}, ..., KLD_{r_N}\}$ and a new BC is trained with the new sets $\mathcal{D}$ and $\tilde{\mathcal{D}}$. This process is repeated until the classification error ε is null or higher than ε_p, the classification error at previous iteration. The threshold ε_p is set experimentally using cross-validation.

Algorithm 5.1: Defect detection learning phase.

Data: $\mathcal{B} = \{B_1, ..., B_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{B}_1, ..., \tilde{B}_m\}$.
Result: Set of reference blocks: $\mathcal{R}$, Bayes classifier: BC.
1 Generate RCT-MoGG signature for each block in $\mathcal{B}$ and $\tilde{\mathcal{B}}$

$\mathcal{R} \leftarrow B_1$; $\varepsilon \leftarrow \infty$; $\mathcal{C} \leftarrow \emptyset$; $N \leftarrow 0$;
repeat
1- $N \leftarrow N + 1$;
2- $\varepsilon_p \leftarrow \varepsilon$;
3- Update the set $\mathcal{R}$ from $\mathcal{C}$;
4- Calculate the KLDs $\{d_1, ..., d_n\}$ and $\{\tilde{d}_1, ..., \tilde{d}_m\}$;
5- Train a Bayes classifier on the KLDs;
6- Classify the ERUs and calculate the error ε;
7- Update $\mathcal{C}$ the set of false defect detections;
until ($\varepsilon = 0$ OR $\varepsilon \geq \varepsilon_p$)

5.4.4.2 The inspection phase

The inspection phase is performed on a newly-seen textile image of the same type as the one used in the learning phase. The steps of the inspection process are given in the script of Algorithm 6.1. First, the input textile image is decomposed into blocks with the same dimension as the training ones. Each block is then classified as containing defects or defect-free using the reference set $\mathcal{R}$ and the trained Bayes classifier BC. Note that in order to achieve very precise localisation of defects, image subdivision into blocks may be performed with overlapping.

Algorithm 5.2: Inspection for defect detection.

Data: Input image I, $\mathcal{R} = \{B_{r_1}, ..., B_{r_N}\}$ and the Bayes classifier BC.
Result: Image with blocks classified.
Decompose the image into blocks;
Generate RCT-MoGG signature for each block;
for (each block B_k) **do**
1- Calculate the KLDs $\{KLD_{r_1}, ..., KLD_{r_N}\}$;
2- Choose $d_k = \min\{KLD_{r_1}, ..., KLD_{r_N}\}$;
3- Use BC to classify B_k.
end for

5.4.4.3 Enhancing defect localisation

A straightforward idea to enhance the defect localisation is to reduce the block size. However, this cannot be achieved by our method since a minimal size of data is required to have stable RCT-MoGG signatures [138]. In fact, the RCT cannot be calculated for images under a minimal size of 64×64 [154]. In addition, significance of statistical parameters for the MoGG models are dependent of the number of data [174, 178].

To enhance the precision of defect localisation without affecting the quality of RCT-MoGG signatures, we devise a new procedure based on subdividing the image into blocks using an appropriate overlapping as illustrated in Fig. (5.7).

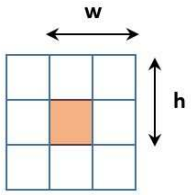


Figure 5.7: Illustration of block overlapping for enhancing defect localisation. In this example, the resolution of defect detection is half the size of an ERU.

This allows for each block to cast one vote for all its contained pixels (1 for defect and 0 for non-defect). Then, the label having the highes t proportion at a given pixel is assigned for its final classification. In the given example, the original block size is $\mathbf{h} \times \mathbf{w}$ and the desired overlapping is 1/2 (half) block. This enables to have a decision at block size $\mathbf{h}/2 \times \mathbf{w}/2$ for the middle colored block. This is achieved by combining the votes from all the original-sized blocks that contain the shaded middle block.

5.5 Experimental results

We have conducted experiments on the popular TILDA database [140], (see Fig. 5.8).

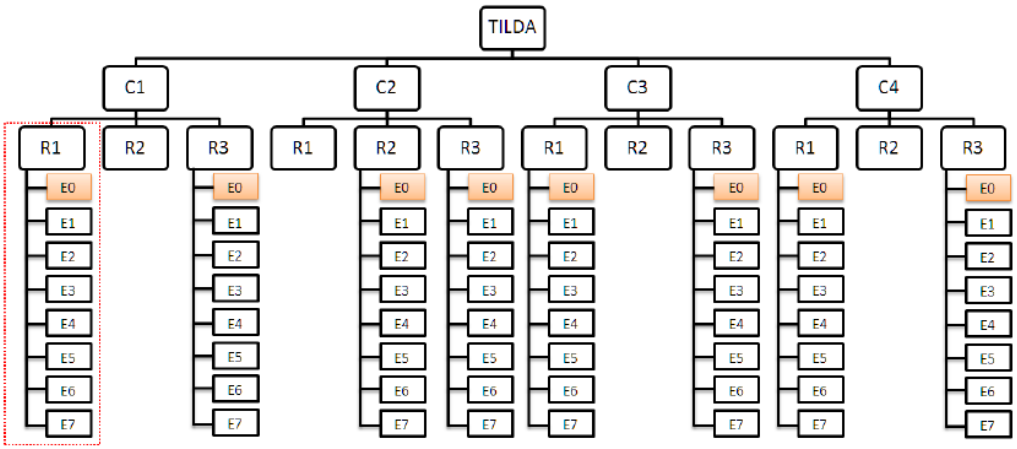


Figure 5.8: TILDA database architecture.

This database consists of eight types of fabric which are stored in four class directories {C1, C2, C3, C4}, and each class directory contains two sub-directories. Therefore, each sub-directory contains one fabric type images which are partitioned into 8 sub-directories containing each fifty texture images. The first sub-directory named 'E0' contains defect-free images, while the other sub-directories ('E1' ... 'E7') contain defective images.

5.5.1 Parameter setting

In the group $p1$ of the dataset, most images contain *plain* and *twill* fabric with texture constituted of very small homogenous patterns. Therefore, we have fixed a block size to 64×64 pixels to allow a reliable estimation of the statistical signatures. In addition, to achieve a good defect localisation, the blocks are spatially overlapped as proposed in Section 5.4.4.3. For fabrics in the non-$p1$ groups, the block size is automatically calculated by our proposed method in Section 5.4.1. For our study, we have used three homogenous fabric types: C1R1, C2R3, C2R3 in the $p1$ group and two periodic fabrics: C3R1 and C3R3 in the non-$p1$ groups. Note that C1R1 is a uniform plain fabric, C2R2 and C2R3 are *twill* and *plain* weave fabrics, respectively. For the non-$p1$ periodic fabrics, C3R1 and C3R3 are gingham ($p4m$ group) and striped fabrics ($p2$ group), respectively, which have either vertical, horizontal or oblique orientations.

5.5.2 Performance evaluation metrics

5.5.2.1 Image-level performance metrics

To be consistent with other methods in the literature, we have first evaluated the performance of our detection algorithm at the image level by calculating detection rates D_R, false alarm rates F_R, detection success rates (also known as detection accuracy) D_{ACC}. These metrics measure the accuracy of tagging images as containing defects or defect-free. These rates are defined as follows:

$$D_R = \frac{TP}{N_{\text{defect}}} \times 100\%, \tag{5.6}$$

$$F_R = \frac{FP}{N_{\text{defect-free}}} \times 100\%, \tag{5.7}$$

$$D_{ACC} = \frac{TP + TN}{TP + FN + TN + FP} \times 100\%, \tag{5.8}$$

where $N_{\text{defect-free}}$ and N_{defect} designate the total numbers of defect-free and defective images, respectively. Table 5.1 gives the definitions of *true positive* (TP), *false positive* (FP), *true negative* (TN) and *false negative* (FN) in the context of defect detection.

	Actually defective	Actually defect-free
Detected as defective	True positive (TP)	False positive(FP)
Detected as defect-free	False negative (FN)	True negative (TN)

Table 5.1: Definition of true positive (TP), false positive (FP), true negative (TN), false negative (FN) in defect detection.

5.5.2.2 Local-level performance metrics

For most of existing methods for defect detection, performance is measured at the image level (i.e., whether an image contains defaults or not). However, such an approach lacks the information about defect localisation which can be important for accessing the accuracy of algorithms. Indeed, even if an image is classified as containing defects by an algorithm, it does not necessarily mean that the right defect is detected. Therefore, we should establish new metrics that can reflect the strength of any method in terms of defect localisation.

In order to evaluate the performance of defect localisation, we propose local metrics calculated at the block level of the image, namely: *Local Precision* (P_L) and *Local Recall* (R_L), often called *hit rate*. We use also *local accuracy* (ACC_L) that is calculated using local *true positive* (TP_L), *true negative* (TN_L), *false positive* (FP_L) and *false negative*(FN_L) values. TP_L is the number of defective blocks identified as such, TN_L is the number of defect-free blocks identified as such, FP_L is the number of defect-free blocks identified as defective and FN_L is the number of defective blocks identified as defect-free. We define the local *precision*, *recall* and *accuracy* metrics as:

$$P_L = \frac{TP_L}{TP_L + FP_L} \times 100\%, \tag{5.9}$$

$$R_L = \frac{TP_L}{TP_L + FN_L} \times 100\%, \tag{5.10}$$

$$ACC_L = \frac{2 \times P_L \times R_L}{P_L + R_L} \times 100\%. \tag{5.11}$$

where ACC_L is the harmonic mean of P_L and R_L (balanced mean between precision and recall), which will be used throughout the experiments to assess the merit of the compared methods.

5.5.3 Results and discussion

5.5.3.1 Image-level performance results

We have compared the performance of our method with the following defect detection methods: *independent component analysis* (ICA) [180], *local binary patterns* (LBP) [181]

and *slope difference distribution* (SDD) [184]. Note that since SDD is a segmentation method based on histogram thresholding [160], it can be applied only for the $p1$ group. For non-$p1$ groups, this method gives arbitrary results for defect detection.

Fig. (5.9) shows some results obtained for images in the $p1$ group using ICA, LBP, SDD and RCT-MoGG with different levels of block overlapping, respectively. We can see that most of methods perform generally well, but the ICA method has failed for the illumination (lighting) defect. Note also that for some images containing small defects (column 5), the LBP method has detected almost the entire image as defective. This is due to the fact that LBP use one experimental threshold which can produce several false positives. In addition, the LBP is sensitive to small changes in the structure of the texture patterns. It is worth pointing that the SDD method has proven some accuracy for localising some defects. However, because the method is based on histogram thresholding [185], it can generate several false positives in case of non-uniform illumination, for example (see columns 4 and 5). Finally, we can see that our approach has yielded good performance and prevents most of the false alarms. In addition, by applying the block overlapping procedure, our method provides a very good accuracy for localising the different defects. Fig. (5.10) shows some examples comparing our algorithm to ICA and LBP on images of non $p1$ groups (striped fabrics). We can note that our algorithm provides better accuracy for localising the defects than the other methods. Finally, some detection results for both $p1$ and non-$p1$ groups are shown in Fig. (5.11) where we can observe that our algorithm has yielded very accurate localisation of defects for all types of fabrics.

For quantitative evaluation, Table 5.2 shows obtained metric values for the homogenous fabrics in the $p1$ group and Table 5.3 for fabrics of non-$p1$ groups in the TILDA dataset [140]. For this type of fabric, we included also the provided results of the recent method SSOCBS [191]. From Table 5.2, we note that our detection rate, ranging from 90% to 100%, is slightly higher than the three compared methods (ICA, SSOCBS and LBP). In addition, the detection accuracy of our method on the fabrics C1R1, C1R3, C2R2, C2R3,

which are 98%, 96.4%, 95% and 91.2%, respectively, are clearly higher than those of the ICA and LBP methods. Note that SSOCBS has not provided figures about the accuracy measure. From Table 5.3, we can observe that the detection rates of our method using fabrics in the non-$p1$ group C3 is 100%. The LBP method exhibits the same result. For these types of fabrics, the ICA almost failed with detection rates ranging from 5 to 30% only. We can also observe for the C3 class in Table 5.3 that the accuracy of our method is higher on average, followed by LBP method, while ICA has the worst performance for this measure.

The most important observation in the above results is the performance of the false alarm rates (FR). Generally, our method generates a lesser number of false alarms than the other methods with both $p1$ and non $p1$ groups (see C1 and C2 in Table 5.2 and C3 in Table 5.3). This means that our method has a good capability to accurately recognize an actual defect-free image as defect-free. This makes our method more effective since this capability is an attractive property required in the textile industry in order to prevent defect-free rejection. This fact is more emphasized in fabrics of the non-$p1$ group, with 0% to 9% false alarms as shown in Table 5.3. The ICA method is very sensitive in the sense that there are acceptable imperfections (small holes and stains) in texture structure that are not regarded as defects but detected as defective, which causes an increase in false alarm rates for some fabrics. This problem is due also to thresholding where small imperfections can disrupt correct detection of defect-free fabrics. Although LBP is more robust than the ICA and is a supervised approach, it suffers from the same problem that it is also based on thresholding.

		C1		C2	
		R1	R3	R2	R3
ICA	D_R (%)	98.5	98	94.5	96
	F_R (%)	**100**	**30**	**52**	**84**
	D_{ACC} (%)	78.8	92.4	85.2	80
LBP	D_R (%)	92.5	94	100	100
	F_R (%)	**36**	**20**	**100**	**100**
	D_{ACC} (%)	87	91	80	80
SSOCBS[191]	D_R (%)	98.5	97.5	97	86
	F_R (%)	-	-	-	-
	D_{ACC} (%)	-	-	-	-
OUR algorithm	D_R (%)	99.5	97.5	96	92
	F_R (%)	**4**	**6**	**10**	**12**
	D_{ACC} (%)	98	96.4	95	91.2

Table 5.2: Comparing the detection rate (D_R), false alarm rate (F_R) and the accuracy (D_{ACC}) of our algorithm versus similar algorithms based on fifty fabric images in each of the five sub-directory (E0, ..., E4).

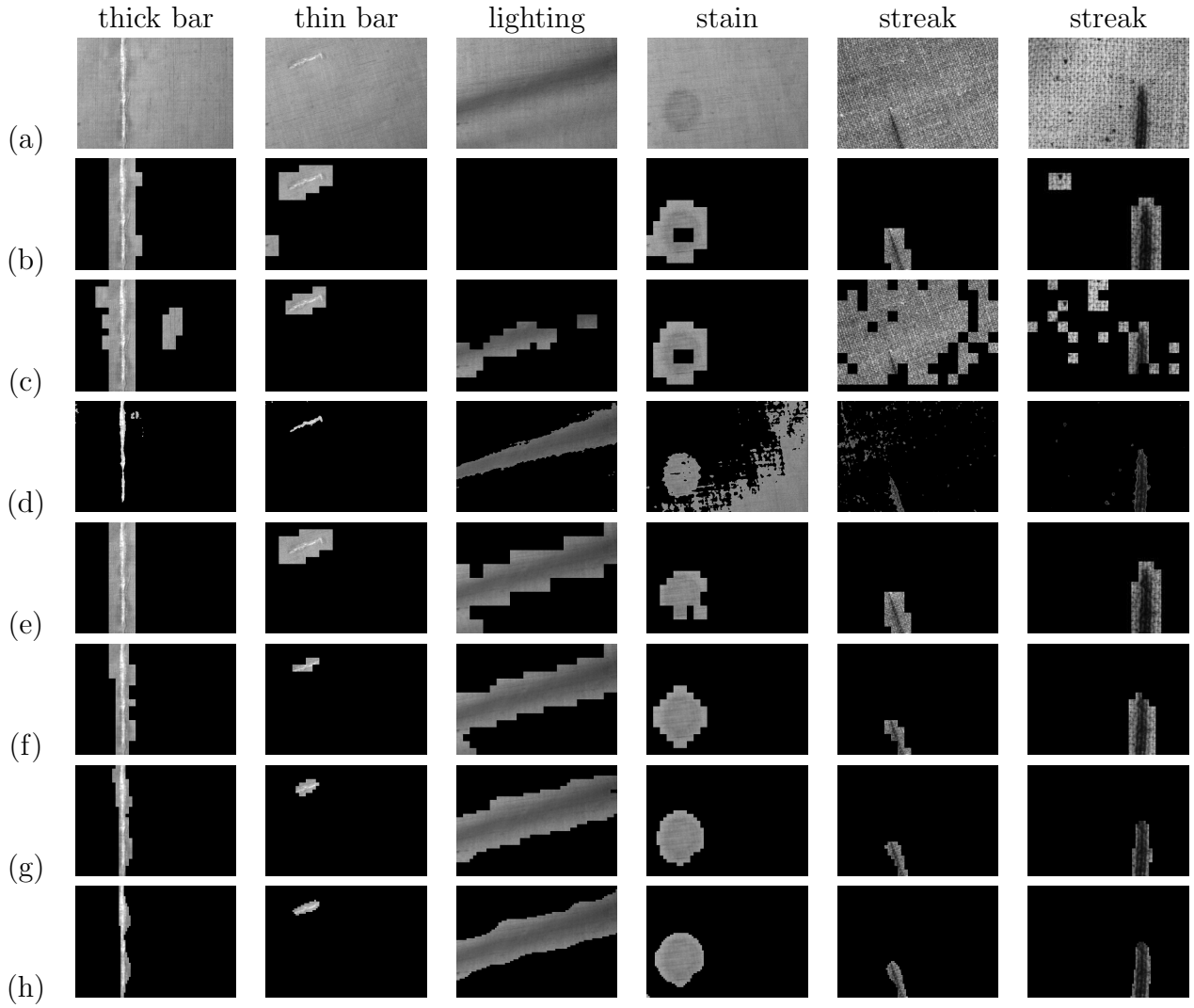


Figure 5.9: Comparison of detection results for homogenous fabrics in $p1$ group from classes C1R1 (columns 1 to 4) C2R2 (column 5) and C2R3 (column 6). (a) represents original fabric images, (b), (c), (d) and (e) represent results obtained using ICA, LBP, SDD and RCT-MoGG without overlapping. (f), (g) and (h) represent results using RCT-MoGG with 1/2, 3/4 and 7/8 block overlapping, respectively.

		C3				
		R1		R3		
		Straight	Oblique	Vertical	Horizontal	Oblique
Our algorithm	D_R (%)	100 (103)	100 (38)	100 (16)	100 (16)	100 (16)
	F_R (%)	**9(11)**	**10 (10)**	**0 (8)**	**0 (8)**	**0 (8)**
	D_{ACC} (%)	99.5 (114)	96 (48)	100 (24)	100 (24)	100 (24)
ICA	D_R (%)	5 (103)	5 (38)	18 (16)	30 (16)	4 (16)
	F_R (%)	9 (11)	**20 (10)**	**50 (8)**	**88 (8)**	**87 (8)**
	D_{ACC} (%)	14 (114)	20 (48)	38 (24)	25 (24)	8 (24)
LBP	D_R (%)	100 (103)	100 (38)	100 (16)	100 (16)	100 (16)
	F_R (%)	**55 (11)**	**100 (10)**	**75 (8)**	**40 (8)**	**88 (8)**
	D_{ACC} (%)	95 (114)	79 (48)	100 (24)	100 (24)	100 (24)

Table 5.3: Comparing the detection rate (D_R), false alarm rate (F_R) and the accuracy (D_{ACC}) of our algorithm with similar ones based on fabric images in each of the five sub-directories (E0, ..., E4). The number of fabrics used for detection are put between brackets. The absence of brackets means all fabric images with defects were used.

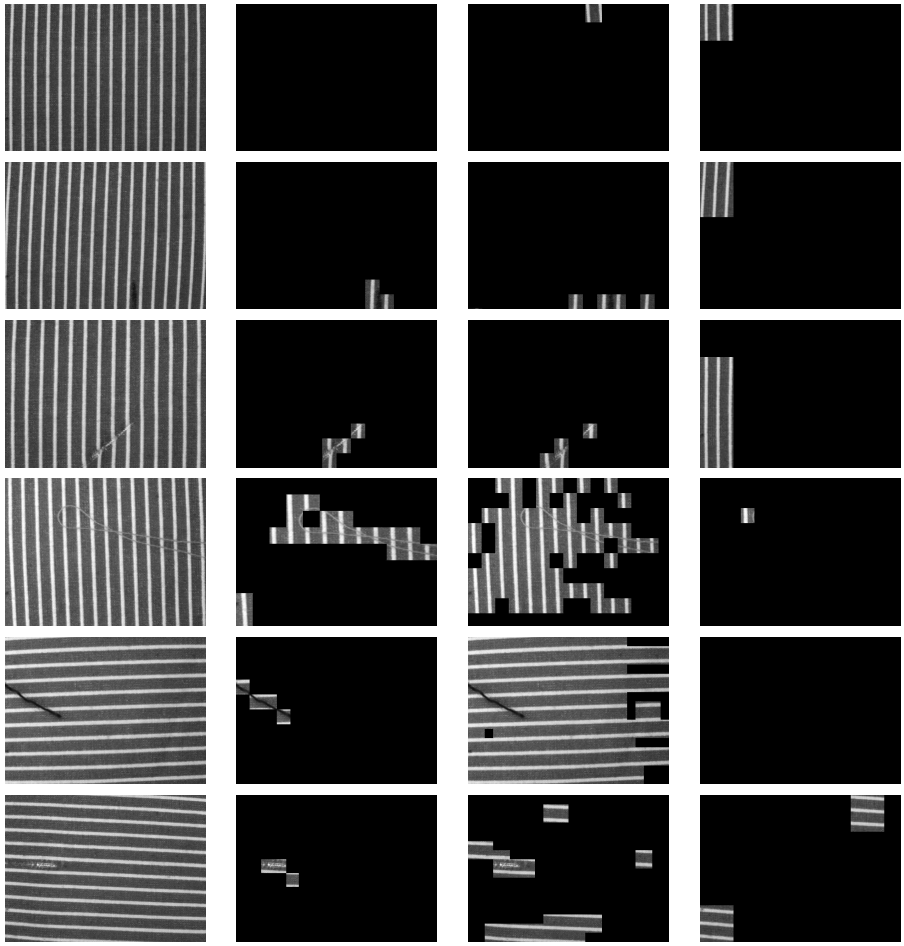


Figure 5.10: Comparison of some detection results for striped fabrics.

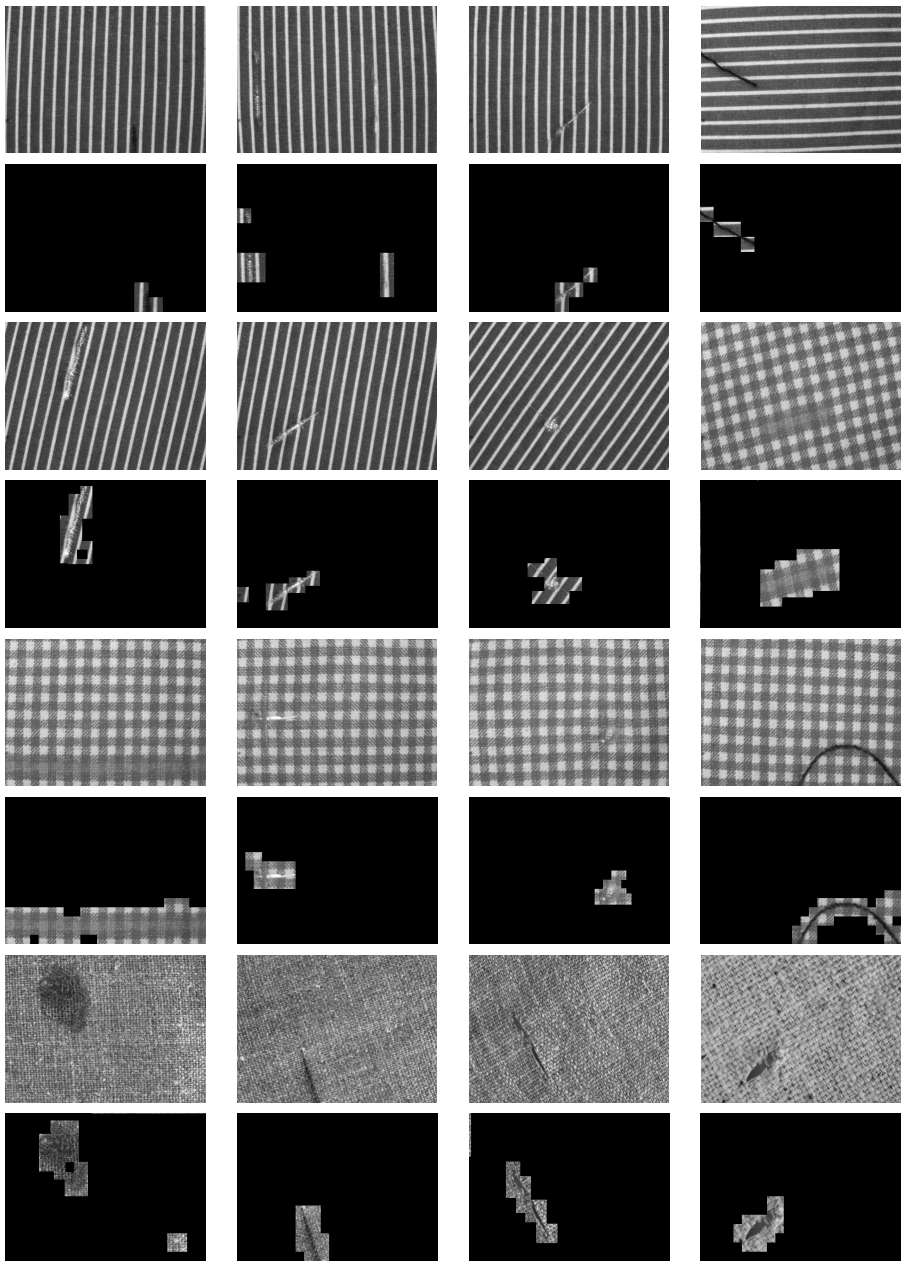


Figure 5.11: Examples of defect detection obtained using our approach on fabric images in the TILDA dataset. The original images are shown in rows 1, 3, 5 and 7 and are as follows: C3R1 (images 1 to 3), C3R3 (images 4 to 7), C2R2 (images 8 to 12) and C2R3 (images 13 to 16). Below each image, we show the detected defects.

5.5.3.2 Local-level performance results

Performance of the proposed method is evaluated in terms of the metrics: local precision (P_L), local recall (R_L) and local accuracy (ACC_L), respectively. We have also built a ground truth where each block of the tested images is tagged as defective or not defective. Using the same numbers of images as in Table 5.2 and Table 5.3, respectively, we have compared the performance of our method with the other methods (ICA and LBP).

The average values of the local metrics calculated for $p1$ and non-$p1$ groups for each method are shown in Tables 5.4 and Table 5.5, respectively. Note that small values of R_L indicate that there are many false negatives, whereas small values of P_L indicate the presence of several false positives generated by an algorithm. Compared to the other methods, our algorithm is more effective in defect localisation where obtained results with regard to he metric ACC_L where values range from (90.6% to 98.3%) for the $p1$ group (see Table 5.4) and (97.8% to 99.2%) for non-$p1$ group (see Table 5.5). Compared to our method, the obtained values of ACC_L by the LBP method, for example, range from 4.1% and 6.4% in fabric C2R2 and C2R3 of the $p1$ group. This clearly indicates that the 100% detection rate (D_R) in Table 5.2 do not actually reflect the effectiveness of this method for defect localisation in patterned fabrics. Also, the ICA method suffers from missing large parts of the defects in the non-$p1$ group of fabrics (C3R1 and C3R3), where ACC_L has drastically dropped.

		C1		C2	
		R1	R3	R2	R3
ICA	P_L (%)	86	85.2	86	78.1
	R_L (%)	90	88.2	87.1	80
	ACC_L (%)	**88**	**86.7**	**86.5**	**79**
LBP	P_L (%)	91.2	93.7	2.1	3.8
	R_L (%)	80	77	100	100
	ACC_L (%)	**85.2**	**84.5**	**4.1**	**6.4**
OUR algorithm	P_L (%)	97.1	95.5	95.8	90
	R_L (%)	99.5	97.4	95.5	91.2
	ACC_L (%)	**98.3**	**96.4**	**95.7**	**90.6**

Table 5.4: Comparing the local precision (P_L), recall (R_L) and accuracy ACC_L of our algorithm versus similar algorithms based on fifty fabric images in each of the five sub-directory (E0, ..., E4).

		C3				
		R1		R3		
		Straight	Oblique	Vertical	Horizontal	Oblique
Our algorithm	P_L (%)	99	98.6	99.2	99	97.5
	R_L (%)	99.4	98.8	99	99	98.2
	ACC_L (%)	**99.2**	**98.7**	**99.1**	**99**	**97.8**
ICA	P_L (%)	4.5	5.6	7.8	9.5	15.2
	R_L (%)	1.1	1.5	1	1	1.1
	ACC_L (%)	**1.8**	**2.4**	**1.7**	**1.8**	**2.1**
LBP	P_L (%)	40.2	25.4	7.1	7.5	5.8
	R_L (%)	85	80.4	77.2	78.6	79.1
	ACC_L (%)	**54.6**	**38.6**	**13**	**13.7**	**10.8**

Table 5.5: Comparing the local precision (P_L), recall (R_L) and and accuracy ACC_L of methods based on fabric images in each of the five sub-directories (E0, ..., E4). The number of test images for these types of fabrics is the same as the specified in Table 5.3.

5.6 Conclusions

We have proposed a new algorithm that has capability to detect and locate defects in most of the fabric groups. The algorithm uses supervised learning to discriminate between defect-free fabrics from defective ones, based on RCT-MoGG signatures. Experiments on several fabric images have shown that our algorithm yields to good detection rates and very low false alarms. Compared to other methods, the proposed algorithm has shown better performance for *plain*, *plain*, *twill* and *weave* fabrics in the $p1$ group, as well as for non-$p1$ groups such as *gingham* fabric ($p4m$) and striped fabric ($p2$) with vertical, horizontal and/or oblique orientations.

Chapter 6

Mixture of multivariate generalized Gaussian with applications

6.1 Motivation

In this chapter, our second contribution is developed. In the pursuit of the best texture representation, we present a unified statistical model for multivariate and multi-modal sub-band coefficient distribution in multi-resolution transforms of color-texture images. This model, based on the formalism of finite mixtures of multivariate generalized Gaussians (MoMGG), enables an accurate description of vectorial coefficients in multi-scale image decomposition. The MoMGG model is proposed using an expectation-maximization algorithm that combines a numerical estimation of parameters via newton-raphson method. The MoMGG not only enables to model inter-band correlation between sub-bands of different scales and orientations but also between intra-band correlation between adjacent locations at different orientations and scales which gives it the capability of better describ-

Figure 6.1: Example image containing of arbitrarily shaped components of random colored textures.

ing the spatial layout of texture patterns. It enables also to seamlessly combine different transforms (e.g., contourlets, wavelets) to build a richer and more representative texture signature for texture search and classification. We successfully applied our approach to texture retrieval and reconstruction where promising results have been obtained comparatively to state-of-art approaches using standard datasets. Fig. 6.1 shows an image of multicomponents of colored textures.

6.2 Introduction

Given the importance visual texture analysis and understanding, texture characterisation has received much scientific attention from vision science, computer vision & graphics, signal processing and arts [33]. The texture analysis field can be structured into three levels of investigation: 1) low-level analysis concerns the representation and modelling of texture, 2) the mid-level analysis which embraces several operations, including discrimination, segmentation and synthesis [4], and 3) high-level analysis which is characterized by domain applications such as remote sensing, medical imaging, inspection and content based image

retrieval [138, 192, 213]. Usually the first level focusing on texture representation, i.e., the extraction of features describing texture information [194] - is basic step for the two other levels and is at the core of texture analysis.

Several methods of textural representation and modelling have been proposed in the past. The point is to find a powerful texture representations lies in balancing of two competing goals: high quality representation and high efficiency. High quality requires texture representations to account for intra-class variations due to changes in illumination, rotation, scale and blur while ensuring distinctiveness between different classes with close appearance. High Efficiency requires methods to be scalable to large amounts of data while being highly compact to deploy in resource-limited platforms such as embedded and hand-held devices. Tuceryan and Jain [4] identified three major methods for texture representation: structural [198, 197], spectral [200, 17, 199, 201], statistical methods [18, 194, 204]. While each method has its advantages and limitations, their combination has been proven to be powerful for building meaningful texture representations. For example, statistical distribution of coefficients in the multi-scale wavelet transform can reveal some properties of first and second order statistics of the image [138, 18]. Therefore, several parametric models have been proposed along with multi-scale transforms for modelling texture images. Parametric models aim at matching the appearance of natural textures by examining artificial textures synthesized by the model [22]. Pioneering works in this regard used the parametric class of Gaussian scale mixtures which allow to generate heavy-tailed distributions by convolving the normal with a uniform distribution [21]. Then, the generalized Gaussian (GGD), Gamma, Student, Pearson Type VII, uniform power, generalized-t [21] have been used to model the distribution of wavelets. The GGD has been particulary explored for its ability to fit platykurtic and leptokurtic distributions [138, 23, 24]. However, since these models are tuned only to univariate representations, they are not able to capture correlations between intra-band and inter-band coefficients.

Multi-band texture modeling of texture features has been investigated in several past works. For example, Markov random fields have been used to model joint statistics of pairs of complex wavelet coefficients at adjacent spatial locations, orientations and scales in texture [22]. This statistical model yields realistic appearance matches for texture analysis and synthesis, and has had a broad impact on the field of human texture perception [210, 211, 33]. In [206], joint modeling of neighboring wavelet coefficients using product of centred Gaussians has been proposed. Although the model has been proven suitable for image denoising, it is not easily applicable for texture image discrimination. In [56], joint alpha-stable sub-Gaussian distribution is used to model wavelet coefficient distributions. In order to better capture the texture statistics, a computationally complex Gaussianization step is required. Other methods to model correlation use spherical invariant random vectors [28] or copulas [29]. [35] used a Gaussian copula-based multivariate model to capture dependencies between complex wavelet coefficients for texture image retrieval. However, the model does not capture multi-modal and is zero-centered data. The full covariance in color texture has been addressed by [26, 27] using multivariate generalized Gaussian density (MGG) to model color wavelet coefficients. However, *multimodality* has not been addressed like [138, 109, 208]. On the other hand, most of past methods focused on describing the distribution for a single transform. To the best of our knowledge, no method has tried to combine coefficients from different multi-resolution transforms in order to increase the accuracy of texture representation. Strictly speaking, the goal is to find a flexible, yet accurate parametric model for describing the joint statistics of arbitrary sub-band combinations while offering a handy way to compare texture images in high-level applications.

Recently, deep learning methods using *Convolutional Neural Networks* (CNNs) have proposed for texture representation [30]. Briefly, weights of filters are learnt using labeled data and convolved with input images to create a spatial feature map of activations, sim-

ilar to a traditional bank of Gabor filters. These methods have the ability to leverage large datasets to learn high quality features. However, their training requires estimating millions of parameters and a very large number of annotated images which requires a huge computational power, an issue which rather constrains their applicability problems with very limited resources and training data. For example, [31] have used CNNs for texture synthesis exploiting several layers of the VGGNet [32]. In [33], the authors compared features encoded by a deep CNN (VGGNet) with the simpler model [22] for texture synthesis and concluded that the former performed slightly better than latter, but no model can synthesize textures that humans cannot discriminate from natural textures.

In the goal of improving the accuracy of parametric models to match the appearance of natural textures, we propose a new semi-parametric statistical model for multivariate and multi-scale decomposition of color texture images. We formulate our model using mixtures of multivariate generalized distributions (MoMGG) where all mixture components are endowed with full covariance matrices to account for full inter-band correlation. This model enables to seamlessly combine arbitrary numbers of sub-bands at different inter-band or intra-band coefficients in multi-scale decompositions. Thanks to the flexibility of the MoMGG, we can fit any distribution of the combined coefficients without restrictions to previous work assumptions about unimodal and centered data. The MoMGG is also able to accurately fit heavy-tailed distributions in a multidimensional setting which characterize natural image statistics [22]. Finally, to compare texture images using MoMGG models, we develop an approximate Kullback-Leibler divergence using monte-carlo sampling. Our contributions can be summarized in the following points:

1. Our MoMGG model extends several previous statistical methods for modelling multi-scale decompositions of texture. Indeed, [26, 29] constitute special cases of our model by assuming unimodal and centered distributions of data. Also, [138, 109] are special cases of the MoMGG by having only one-dimensional data. In addition to its ability

to model multi-modal and heavy-tailed distributions, the MoMGG enables accurate joint modeling of arbitrary combinations of intra-band and inter-band texture information with full covariance potential. To estimate the MoMGG parameters, we propose a method based on the expectation-maximization algorithm. Furthermore, to ensure a suitable similarity measurement between MoMGG models, and thus compare images, we propose an approximate Kullback-Leiber divergence by using multidimensional monte-carlo sampling. Beside being computationally efficient.

2. For texture representation, we explore MoMGG modelling using different combination scenarios of multi-scale decomposition sub-bands. In the simple scenario, we combine bands with the same scale and orientation generated from different color channels. In the second scenario, we combine subbands with different orientations at the same scale, then with orientations at different scales. In the third scenario, we model cross-joint correlation between spatial translational of the image. This combination aims at finding a way to better describe the spatial layout of the texture patterns. These different scenarios are incremental in the sense that scenario 3 is an extension of scenario 2 which is in intrinsically an extension of scenario 1. In the final scenario, we combine subbands of different image transforms (e.g., contourlets and wavelets).

3. To evaluate the potential of the MoMGG model for texture representation, we conducted experiments on content based color-texture retrieval (CBIR) and image reconstruction using various standard datasets. We compared our performance with recent methods based on parametric statistical models, all based on compact modelling of multi-scale decomposition coefficients. Qualitative and quantitative evaluation results have shown that our method performs tetter in both CBIR and image reconstruction applications.

The remainder of this chapter is organized as follows. Section 2 presents our MoMGG

model parameter estimation. Section 3 presents the texture modeling and retrieval using MoMGG. Section 4 present derivation of multivariate KLD sampling between MoGGDs. Section 5 presents the experimental results. Finally, Section 6 concludes the chapter and we also provide an outlook toward possible improvements.

6.3 Mixture of multivariate generalized Gaussians (MoMGG)

6.3.1 Formulation of the MGG

The density of a random vector $\mathbf{x}$ following a p-dimensional *multivariate generalized Gaussian* (MGG) distribution (also known as the power exponential distribution) is given as follows [221]:

$$f(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \beta) = r|\boldsymbol{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}\delta(\boldsymbol{x})^{\beta}\right\}, \tag{6.1}$$

where

$$r = \frac{p\Gamma(\frac{p}{2})}{\pi^{\frac{p}{2}}\Gamma\left(1+\frac{p}{2\beta}\right)2^{1+\frac{p}{2\beta}}},$$

with $\delta(\boldsymbol{x}) = (\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$. The parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and β give the location (mean), a positive-define scale matrix and a positive shape parameter, respectively. The parameter β controls the peakedness of the distribution and the heaviness of its tails. When $\beta = 1$ we retrieve the multivariate Gaussian (MG) distribution, in this case $\boldsymbol{\Sigma}$ becomes the covariance matrix. When $\beta = 1/2$, we obtain the multivariate Laplacian (ML) distribution. Furthermore, when $\beta \longrightarrow \infty$ the MGG distribution becomes a multivariate uniform distribution. The covariance and multidimensional kurtosis coefficient for the MGG distribution

are given by [221]:

$$\mathbf{C} = \frac{2^{1/\beta}\Gamma\left(\frac{p+2}{2\beta}\right)}{p\Gamma\left(\frac{p}{2\beta}\right)}\mathbf{\Sigma} \tag{6.2}$$

and

$$\gamma_2 = \frac{p^2\Gamma\left(\frac{p}{2\beta}\right)\Gamma\left(\frac{p+4}{2\beta}\right)}{\Gamma^2\left(\frac{p+2}{2\beta}\right)} - p(p+2), \tag{6.3}$$

γ_2 denotes the multidimensional kurtosis [221]. For $\beta \in (0,1)$, the MGG distribution is a scale mixture of Gaussian distributions [223]. The MGG has been extensively used in the literature [130, 224]. Since the estimation of the MGG distribution parameters can not be obtained in a closed form, especially the covariance over the entire support of the shape parameter $\beta \in (0,\infty)$, some authors resorted to numerical techniques such as the Newton-Raphson method [127]. The authors have focused mainly on the special case where $0 < \beta < 1$.

6.3.2 Mixture of MGGs (MoMGG) and model estimation

It is well known that finite mixture models provide a convenient yet a formal framework for investigating heterogeneity in data, which makes them suitable for clustering and classification. Based on the MGG distribution, a mixture of MGGs (MoMGG) can be defined as:

$$h(\boldsymbol{x}|\Theta) = \sum_{k=1}^{K} \alpha_k f(\boldsymbol{x}|\boldsymbol{\mu}_k, \mathbf{\Sigma}_k, \beta_k), \tag{6.4}$$

where $f(\cdot)$ is the k^{th} component density and Θ denotes the set of parameters. $\boldsymbol{\mu}_k$, $\mathbf{\Sigma}_k$ and β_k denote the mean, scale matrix, and shape parameter, respectively, of the k^{th} component. The parameters $\alpha_1, \ldots, \alpha_K$ are the mixing weights such that $\alpha_k > 0\,(k = 1, ..., K)$

and $\sum_{k=1}^{K} \alpha_k = 1$. Fig. 6.2 presents in (a) and (c) two mixtures of bivariate generalized Gaussians and their respective iso-contour plots in (b) and (d). We can see in (a) that the MGGs are leptokurtic with heavy tails with the same shape parameter $\beta = 0.6$. The model in (c) is composed of a peaked bivariate Laplacian with $\beta = 1/2$ and a flat (uniform) bivariate generalized Gaussian with $\beta = 8$. We note that as $\beta \to \infty$, the MMG converges to a multivariate uniform distribution. The directions of the main axes of the iso-contours centered on their respective means, clearly show correlations in the data.

Herein we develop a mixture of MGGs in which the full covariance over the entire support of the shape parameter $\beta \in (0, \infty)$ is estimated using the expectation-maximization algorithm. Furthermore, our approach combines a numerical estimation of the parameters via the Newton-Raphson method and the *minimum message length* (MML) principle for model selection [138]. Given a sample of data $\mathcal{S} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_N\}$, the likelihood of the MoMGG is defined as:

$$L(\Theta|\mathcal{S}) = \prod_{i=1}^{N} \sum_{k=1}^{K} \alpha_k r_k |\boldsymbol{\Sigma}_k|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \left[\delta_{ik}(\boldsymbol{x}_i)\right]^{\beta_k} \right\}, \tag{6.5}$$

where $\delta_{ik}(\boldsymbol{x}_i) = (\boldsymbol{x}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{x}_k - \boldsymbol{\mu}_k)$ and $r_k = \frac{p\Gamma(\frac{p}{2})}{\pi^{\frac{p}{2}} \Gamma\left(1+\frac{p}{2\beta_k}\right) 2^{1+\frac{p}{2\beta_k}}}$.

The $\mathcal{S}$ is considered incomplete since no information is provided about the component membership of each data instance. The complete data is given by $\mathcal{S}_c = \{(\boldsymbol{x}_1, \boldsymbol{z}_1), \ldots, (\boldsymbol{x}_N, \boldsymbol{z}_N)\}$, where the missing variable $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{iK})^T$ is the mixture component label vector such that $z_{ik} = 1$ if $\boldsymbol{x}_i$ comes from the k^{th} component and 0 otherwise, hence $z_{ik} \in \{0, 1\}$ and $\sum_{k=1}^{K} z_{ik} = 1$. By applying the logarithm, the complete log-likelihood is then given by:

$$\mathcal{L}_c(\Theta) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \log \left[\alpha_k r_k |\boldsymbol{\Sigma}_k|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \left[\delta_{ik}(\boldsymbol{x}_i)\right]^{\beta_k} \right\} \right] \tag{6.6}$$

The estimation of the parameters α_k, β_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, $k = (1, ..., K)$, can not be estimated in a closed form. We use rather the expectation-maximization (EM) algorithm which is

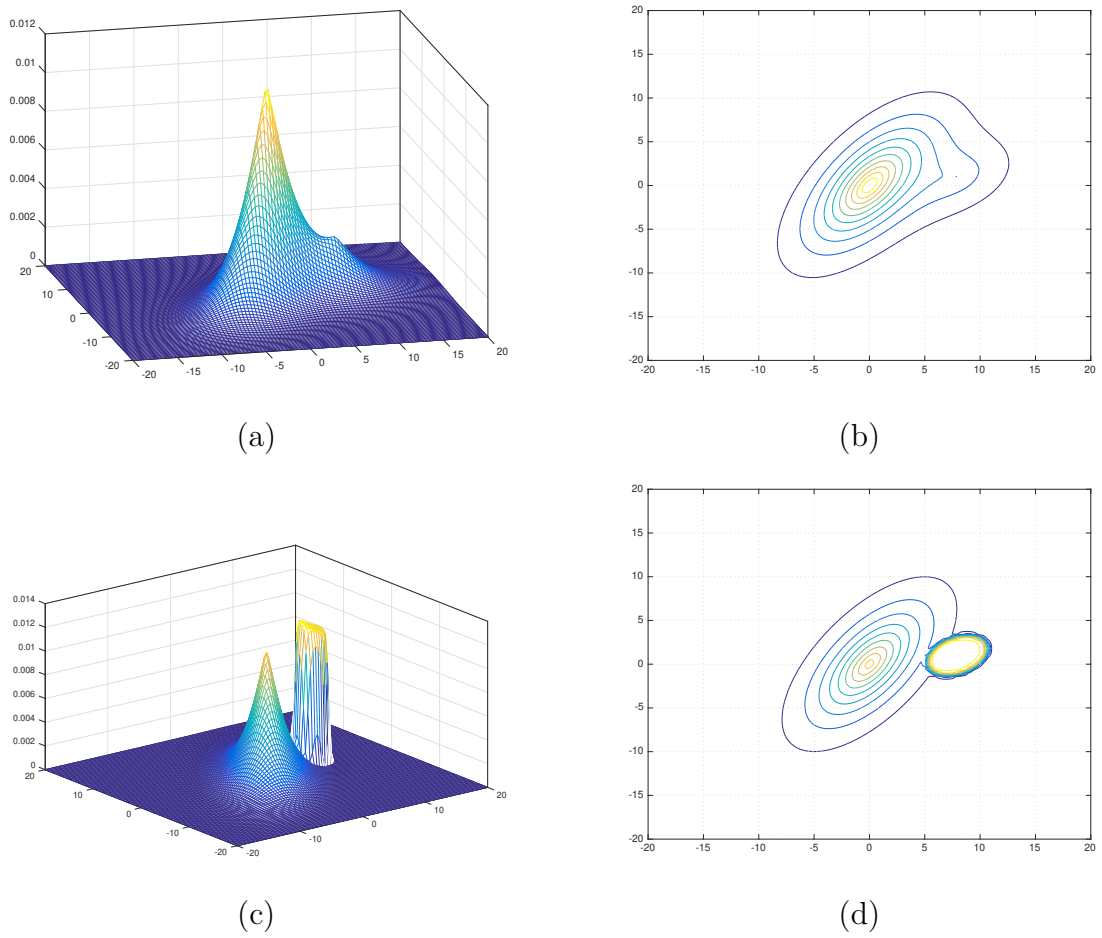


Figure 6.2: Examples of mixtures of bivariate generalized Gaussians: (a) A mixture of two MGGs plotted for the same shape parameter $\beta = 0.6$ and different values of full covariances matrix $\boldsymbol{\Sigma}$ and means $\boldsymbol{\mu}$, (b) represents the isocontours of mixture (a), (c) a mixture of two MGGs, one is a bivariate Laplacian ($\beta = 1/2$) and the other is a bivariate uniform density ($\beta = 8$), with their isocontours in (d).

an iterative method based on the complete-data log-likelihood (6.6). The E-step involves calculating the expected complete-data log-likelihood. We need the expected values given by:

$$\tau_{ik}^{(t)} = \mathbb{E}_{\widehat{\boldsymbol{\Theta}}^{(t)}}[Z_{ik}|\boldsymbol{x}] = \frac{\hat{\alpha}_k f(\boldsymbol{x}_i|\widehat{\boldsymbol{\mu}}_k, \hat{\boldsymbol{\Sigma}}_k, \hat{\beta}_k)}{\sum_{j=1}^{K} \hat{\alpha}_j f(\boldsymbol{x}_i|\widehat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\Sigma}}_j, \hat{\beta}_j)}, \tag{6.7}$$

for $i = (1, \ldots, N)$ and $k = (1, \ldots, K)$. Here $\hat{\alpha}_k$, $\widehat{\boldsymbol{\mu}}_k$, $\hat{\boldsymbol{\Sigma}}_k$, $\hat{\beta}_k$ are the estimation of the parameters obtained from he tth iteration. The M-step on $(t+1)^{\text{th}}$ iteration involves maximization of the expected values of the complete data log-likelihood with respect to the parameters. The update for α_k^{t+1} can simply be obtained as follows:

$$\hat{\alpha}_k^{(t+1)} = \frac{\sum_{i=1}^{N} \tau_{ik}^{(t)}}{N} \tag{6.8}$$

The maximization of the expected log-likelihood with respect to the parameters β_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ can be obtained using the Newton-Raphson method. To find $\boldsymbol{\mu}_k$ update, we calculate the first and second derivatives of the expected complete log-likelihood $\mathcal{Q}(\Theta, \hat{\Theta})$ with respect to $\boldsymbol{\mu}_k$, as follows:

$$\frac{\partial \mathcal{Q}}{\partial \boldsymbol{\mu}_k} = \hat{\beta}_k \sum_{i=1}^{N} \tau_{ik} \delta_{ik}(\boldsymbol{x}_i)^{\hat{\beta}_k - 1} \hat{\boldsymbol{\Sigma}}_k^{-1} m_{ik} \tag{6.9}$$

$$\frac{\partial^2 \mathcal{Q}}{\partial \boldsymbol{\mu}_k^2} = \hat{\beta}_k \sum_{i=1}^{N} \tau_{ik} \left[-\delta_{ik}(\boldsymbol{x}_i)^{\hat{\beta}_k - 1} \hat{\boldsymbol{\Sigma}}_k^{-1} + (\hat{\beta}_k - 1) \delta_{ik}(\boldsymbol{x}_i)^{\hat{\beta}_k - 2} \hat{\boldsymbol{\Sigma}}_k^{-1} m_{ik} \left(-2 \hat{\boldsymbol{\Sigma}}_k^{-1} m_{ik} \right)^T \right] \tag{6.10}$$

where $m_{ik} = (\boldsymbol{x}_i - \hat{\boldsymbol{\mu}}_k)$. Likewise, the derivatives for the update of β_k are given as follows:

$$\frac{\partial \mathcal{Q}}{\partial \beta_k} = \frac{p n_k}{2\hat{\beta}_k^2} \psi(a) + \frac{p n_k \log(2)}{2\hat{\beta}_k^2} - \sum_{i=1}^{N} \frac{\tau_{ik}}{2} \delta_{ik}(\boldsymbol{x}_i)^{\hat{\beta}_k} \log \delta_{ik}(\boldsymbol{x}_i) \tag{6.11}$$

$$\frac{\partial^2 \mathcal{Q}}{\partial \beta_k^2} = \frac{-p n_k}{\hat{\beta}_k^3} \psi(a) - \frac{p^2 n_k}{4\hat{\beta}_k^4} \psi_1(a) - \frac{p n_k \log(2)}{\hat{\beta}_k^3} - \sum_{i=1}^{N} \frac{\tau_{ik}}{2} \delta_{ik}(\boldsymbol{x}_i)^{\hat{\beta}_k} [\log \delta_{ik}(\boldsymbol{x}_i)]^2 \tag{6.12}$$

where $a = 1 + \frac{p}{2\hat{\beta}_k}$, $\psi(\cdot)$ and $\psi_1(\cdot)$ are the digamma and trigamma functions and $\sum_k^K n_k = N$.

We can find estimation of the scale matrix $\boldsymbol{\Sigma}_k$ by ignoring terms not involving $\boldsymbol{\Sigma}_k$ in the complete log-likelihood:

$$\mathcal{Q}(\boldsymbol{\Sigma}_k) = \sum_{i=1}^{N} \sum_{i=k}^{K} \frac{\tau_{ik}}{2} \log |\boldsymbol{\Sigma}_k|^{-1} - \frac{\tau_{ik}}{2} \left(m_{ik}^T \boldsymbol{\Sigma}_k^{-1} m_{ik} \right)^{\beta_k} \tag{6.13}$$

The updates could be obtained by some numerical methods such as fixe point, minorization-maximization framework and Newton-Raphson. Upon taking the derivative of $\mathcal{Q}(\mathbf{\Sigma}_k)$ with respect to Σ^{-1}, we can obtain the following fixed-point update:

$$\frac{\partial \mathcal{Q}}{\partial \Sigma_k} = \sum_{i=1}^{N} \frac{\tau_{ik}}{2}\Sigma_k - \tau_{ik}\beta_k \left(m_{ik}^T \mathbf{\Sigma}_k^{-1} m_{ik}\right)^{\beta_k - 1} m_{ik} m_{ik}^T \tag{6.14}$$

$$\frac{\partial^2 \mathcal{Q}}{\partial {\Sigma_k}^2} = \sum_{i=1}^{N} \beta_k(\beta_k - 1)\left(m_{ik}^T \mathbf{\Sigma}_k^{-1} m_{ik}\right)^{\beta_k - 2} vec\left(m_{ik} \otimes m_{ik}\right) vec\left(\mathbf{\Sigma}_k^{-1} m_{ik} m_{ik}^T \mathbf{\Sigma}_k^{-1}\right)^T \tag{6.15}$$

where $vec(\cdot)$ is a transformation of a matrix into a vector.

The parameter estimation and model selection of the MoMGG model are summarized in the algorithm given below. Given a candidate number$K$ of components, the mixture parameters are estimated iteratively using the expectation-maximization (EM) steps. Note that the convergence of the EM is detected when the distance between the parameters resulting from two successive iterations l and $l+1$ is smaller than a pre-defined threshold ϵ, i.e. $||\Theta^{l+1} - \Theta^l|| < \epsilon$. Note also that the initialization of parameters is performed using mixture of multivariate Gaussian densities. Then, we iterate on mixture of multivariate generalized Gaussians (MoMGG) by readjusting the parameters. The model selection and parameter estimation of the MoGG are achieved in an unsupervised fashion using the minimum message length (MML) principle[138].

6.4 Multivariate texture modelling using MoMGG

6.4.1 Multi-band texture combinations for MoMGG modelling

To take advantage of the MoMGG to model multivariate data, we propose different sceneries of combining arbitrary numbers of bands in multi-scale decomposition of color texture. Contrary to past methods describing only the marginal distributions of different bands

Algorithm 6.1: Parameter estimation and model selection for the MoMGG.

Data: Input data : $\mathcal{S} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_N\}$.
  for $K = K_{min}$ to K_{max} **do**
    Initialize $\hat{\beta}_k$, $\hat{\boldsymbol{\Sigma}}_k$, $\hat{\boldsymbol{\mu}}_k$, $\hat{\alpha}_k$, $(k = 1, ..., K)$.

    repeat
      Update $\hat{\alpha}_k$ using (6.8);
      Update $\hat{\boldsymbol{\mu}}_k$ using (6.9) and (6.10).
      Update $\hat{\beta}_k$ using (6.11) and (6.12).
      Update $\hat{\boldsymbol{\Sigma}}_k$ using (6.14) and (6.15).
    until (Convergence)
    Calculate the associated minimum message length (MML) criterion;
  end for
  Select the optimal model K^* such that $K^* = argmin_K MessLength(K)$.

[138], our statistical model provides a powerful way to describe jointly the distribution of multiple bands. Analyzing the joint statistics of the different band aims at extracting valuable information about the texture structure at different color bands as well as between different orientations and scales.

As presented in the introduction, we build different types of combinations of texture channels in the multi-scale decomposition of images (as shown in Fig. 6.3). In the simplest scenario, we combine bands with the same scale and orientation generated from different color channels. The second scenario combines subbands with different orientations at the same scale, then with orientations at different scales. The third scenario combines spatially-translated bands of the image in order to capture the texture layout of the texture patterns. These two scenarios are similar to the ones proposed in the seminal work [22]. The final scenario consists of combining different decompositions transforms such as wavelets and contourlets [109] in order to take advantage of the benefits of each transform. All these combinations will generate multidimensional vectors that will be combined using the MoMGG model.

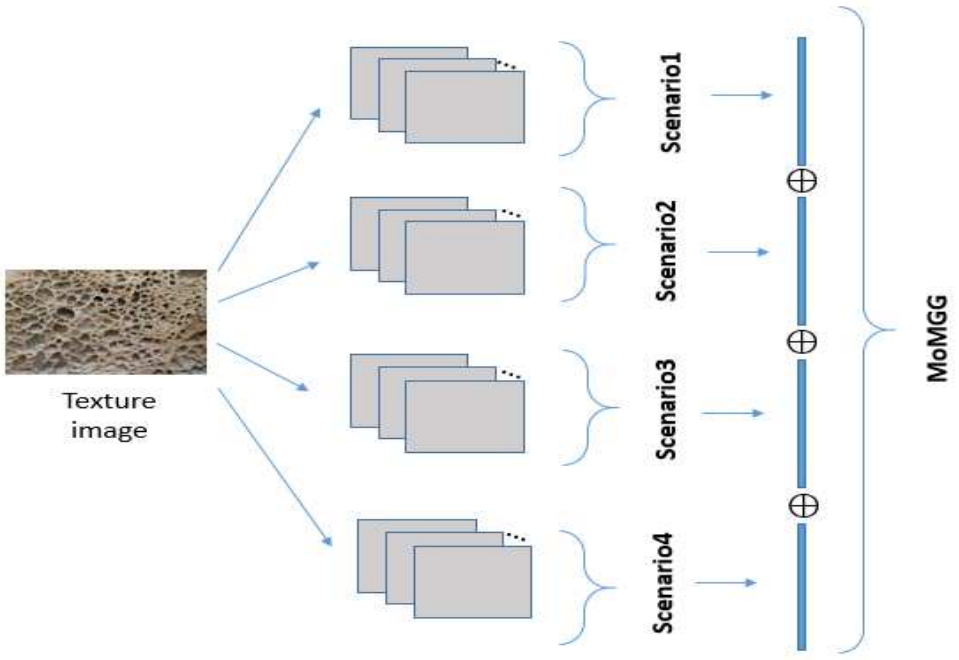


Figure 6.3: Different scenarios of multi-band combination in the MoMGG model.

6.4.2 Similarity measurement between texture images

Since each texture image is represented by a MoMGG model, the similarity between images will be determined by the distance between their representative MoMGG models. In general, let $P(\boldsymbol{x})$ and $Q(\boldsymbol{x})$ be two probability distributions of a multi-dimensional variable $\boldsymbol{x} \in \mathbb{R}^p$. Typically $P(\boldsymbol{x})$ represents the "true" distribution of data (observations), or a precisely calculated theoretical distribution, whereas $Q(\boldsymbol{x})$ represents a model (or approximation) of $P(\boldsymbol{x})$. To measure the difference between two probability distributions over the same variable $\boldsymbol{x}$, there exist a number of similarity measures [212]. Among the most popular measure, the *Kullback-leiber divergence* (KLD) has been extensively used in the literature.

The KLD, which is closely related to relative entropy and mutual information, is a non-symmetric measure of the difference between two probability distributions $P(\boldsymbol{x})$ and

$Q(\boldsymbol{x})$. Specifically, the KLD of $P(\boldsymbol{x})$ from $Q(\boldsymbol{x})$), denoted by $KLD(P,Q)$, is a measure of the information lost when $Q(\boldsymbol{x})$ is used to approximate $P(\boldsymbol{x})$. Its formulation is given by:

$$KLD(P||Q) = \int P(\boldsymbol{x}|\Theta_P) ln \frac{P(\boldsymbol{x}|\Theta_P)}{Q(\boldsymbol{x}|\Theta_Q)} dx \tag{6.16}$$

The KLD from P to Q is not a distance since it is not symmetric and does not satisfy the triangular inequality. Nevertheless, $KLD(P||Q)$ is a non-negative measure and expresses a good similarity measure between two distributions. Note that $KLD(P||Q) \geq 0$ and $KLD(P||Q) = 0$ if and only if $P = Q$. Note that a closed-form KLD between two centered MGG distributions in the case $\beta = 2$ (normal distributions).

$$KLD(\Sigma_1||\Sigma_2) = \frac{1}{2}\left[ln \frac{|\Sigma_2|}{|\Sigma_1|} + tr(\Sigma_2^{-1}\Sigma_1 - p\right] \tag{6.17}$$

with Σ_1 and Σ_2 are respective covariances of P and Q. The closed-form extension to the general case between centered MGGs (where $\beta_1 \neq \beta_2 \neq 1$)) has been obtained in [27] as follows:.

$$\begin{aligned} KLD(\beta_1, \Sigma_1||\beta_2, \Sigma_2) = ln \left[\frac{\Gamma(\frac{1}{\beta_2})}{\Gamma(\frac{1}{\beta_1})} 2^{(\frac{1}{\beta_2} - \frac{1}{\beta_1})} \left(\frac{|\Sigma_2|}{|\Sigma_1|} \right)^{\frac{1}{2}} \frac{\beta_1}{\beta_2} \right] \\ + \left[-\frac{2^{\frac{1}{\beta_1}}}{\beta_1} + 2^{\frac{\beta_1}{\beta_2}} \frac{\Gamma(\frac{\beta_2+1}{\beta_1})}{\Gamma(\frac{1}{\beta_1})} \times \left(\frac{\gamma_1 + \gamma_2}{2} \right)^{\beta_2} 2^{\mathbf{F}\left(\frac{1-\beta_2}{2}, -\frac{\beta_2}{2}; 1; A^2\right)} \right] \end{aligned} \tag{6.18}$$

where $\gamma_i \equiv (\lambda_2^i)^{-1}$, $i = 1, 2$, λ_2^i the eigenvalues of $\Sigma_1^{-1}\Sigma_2$. $A \equiv \frac{\gamma_1 - \gamma_2}{\gamma_1 + \gamma_2} \cdot 2^{\mathbf{F}(.,.;.;.)}$ the hypergeometric function of Gauss which can be tabulated for realistic values of β and $-1 < A < 1$. For $\beta_1 = \beta_2 = 2$ the hypergeometric function becomes 1 and Eq. (6.18) is reduced to Eq. (6.17). However, the above formulation is not generalizable to non-centered MGGs, let alone for calculating the KLD between MoMGGs.

When these distributions are multimodal, a closed-form solution is intractable even for

univariate data [138]. One way to circumventing this issue is by designing a fast way to approximate the KLD between two unconstrained mixture models. The approximation is based on Monte-Carlo sampling methods [109] where given two MoMGG models $P(\boldsymbol{x}) = \sum_{i=1}^{K} \pi_i P(\boldsymbol{x}|\theta_i)$ and $Q(\boldsymbol{x}) = \sum_{j=1}^{M} \omega_j q(\boldsymbol{x}|\alpha_j)$, the approximate KLD between these models can be defined as follows:

$$KLD_{mc}(P||Q) = \frac{1}{n}\sum_{i=1}^{n} \log\frac{P(\boldsymbol{x}_i)}{Q(\boldsymbol{x}_i)} \approx_{n\to\infty} KLD(P||Q) \tag{6.19}$$

The approximation method by sampling aims to generate a sufficiently large sample $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ drawn independently from P in order to approximate the KLD using the Monte-Carlo integration [139]. The sample in our case is generated using the Metropolis-Hasting (MH) algorithm [139] which is a powerful Markov Chain Monte Carlo method (MCMC) for producing dependent simulations from an arbitrary distribution. In brief, the MH algorithm involves the construction of a Markov sequence of random variables whose equilibrium distribution is the desired posterior distribution. In our case, we used $n = 10^5$ as typical value which gives satisfactory results.

6.5 Experimental setup

To measure the performance of the proposed approach, we conducted a series of experiments using three benchamark texture databases. The first one is ALOT [218], containing 250 different texture classes with 100 images per class. The second one is the 40 textures selected from MIT Vistex [3] database called VisTex(small). The third one is the full VisTex which consists of 167 textures, called VisTex (full). In the different combination scenarios described in Section 6.4.1, we used mainly the Daubechies wavelets for the multi-scale decomposition except for the forth scenario where wavelets are coupled with contourlets. For the third scenario, we applied three horizontal and three vertical shifts to generate spatial layout information about the texture.

The evaluation of our method is done mainly in the context of content-based image retrieval. We carried out a comparative evaluation with several advanced recognition methods on the three databases under the same experimental conditions as the authors, basing on our four modeling scenarios. However, as a first experiment, we tested our approach only on the ALOT database in order to evaluate its performance for different decomposition(scale) levels (1 to 3). All our experiments were performed in the MATLAB environment.

6.5.1 Scale levels evaluation with ALOT database

In this experiment we retain 20 classes corresponding to different themes. Similar to many previous researches in the literature, we have divided the original images 512×512 into 16 non-overlapping sub-images 128×128. Therefore, the number of sub-images used in the evaluation are 400 in each class and 8000 in the whole dataset. We measured the retrieval performance in terms of the average retrieval rate (%) which refers to the fraction of the number of relevant images found among the top-retrieved images. More formally, the retrieval rate for each query is given as: $ARR = RI/R$, where ARR is average retrieval rate, RI is the number of relevant images among top N retrieved images from the dataset and R is the total number of relevant images which is always 16 in our case. A sub-image is considered relevant if it is a part of the same original 512512 image as the query sub-image. The similarity between the query image and a target image is computed thanks to the MCMC-KLD approximation in Eq. (6.19).

To concretely evaluate the retrieval performance of the proposed approach, each of 25 images in each class from the dataset is considered as a query image. So in each class, the retrieval rate is calculated by averaging retrieval rates corresponding to 250 $= 10 \times 25$ randomly-selected queries (10 random queries for each image). Therefore in the whole dataset, the retrieval rate is calculated by averaging retrieval rates correspond-

ing to 5000 = 10 × 25 × 20 randomly-selected queries (10 random queries for each image for each of 20 classes). The next sections show evaluation results where we denote scenario 1 as MoMGG+KLD+CD, where CD denotes color-based descriptor, scenario 2 MoMGG+KLD+CCOS, where CCOS denotes cross-correlation between different orientations and scales, scenario 3 MoMGG+KLD+Sh, where Sh denotes image shift and scenario 4 by MoMGG+KLD+CombT, where CombT indicates transform combinations.

Table 6.1 shows influence of scale level on the average retrieval rate for the top 16 retrieved images by considering the four scenarios, respectively. We can observe that increasing the number of decomposition levels improves generally the retrieval accuracy. However, the first two levels embody most of the discrimination power in the high-pass subbands. The third level improves slightly the retrieval accuracy when combined with the first two levels. On the other hand, keeping low-pass subbands improves the retrieval accuracy leading up to 100%.

Level J	scenario 1	scenario 2	scenario 3	scenario 4
J=1	94	97,5	98,2	94,5
J=2	95	97,7	98,7	96
J=3	96	98,5	99,7	97,1
J=3 + A	97	99,2	99,8	98

Table 6.1: Average retrieval rate (%) in the top 16 images retrieved in 5000 queries in the ALOT dataset. MoMGG+KLD+CD (scenario 1),MoMGG+KLD+CCOS (scenario 2), MoMGG+KLD+Sh (scenario 3) and MoMGG+KLD+CombT (scenario 4).

Fig. 6.4 presents a comparison of retrieval precision as a function of the number of top matches considered. It exhibits the average retrieval rate in the dataset of 5000 queries using different levels of decomposition J and the image approximation for all studied scenarios. From model (scenario) 1 to 3, we observe that the results improve from one model to another. This is due to the incremental process where model 2 improves model 2 and model 3 improves model 2. Note that model 4 is a different model from the three

preceding models since it combines different transforms, namely wavelet and contourlet coefficients.

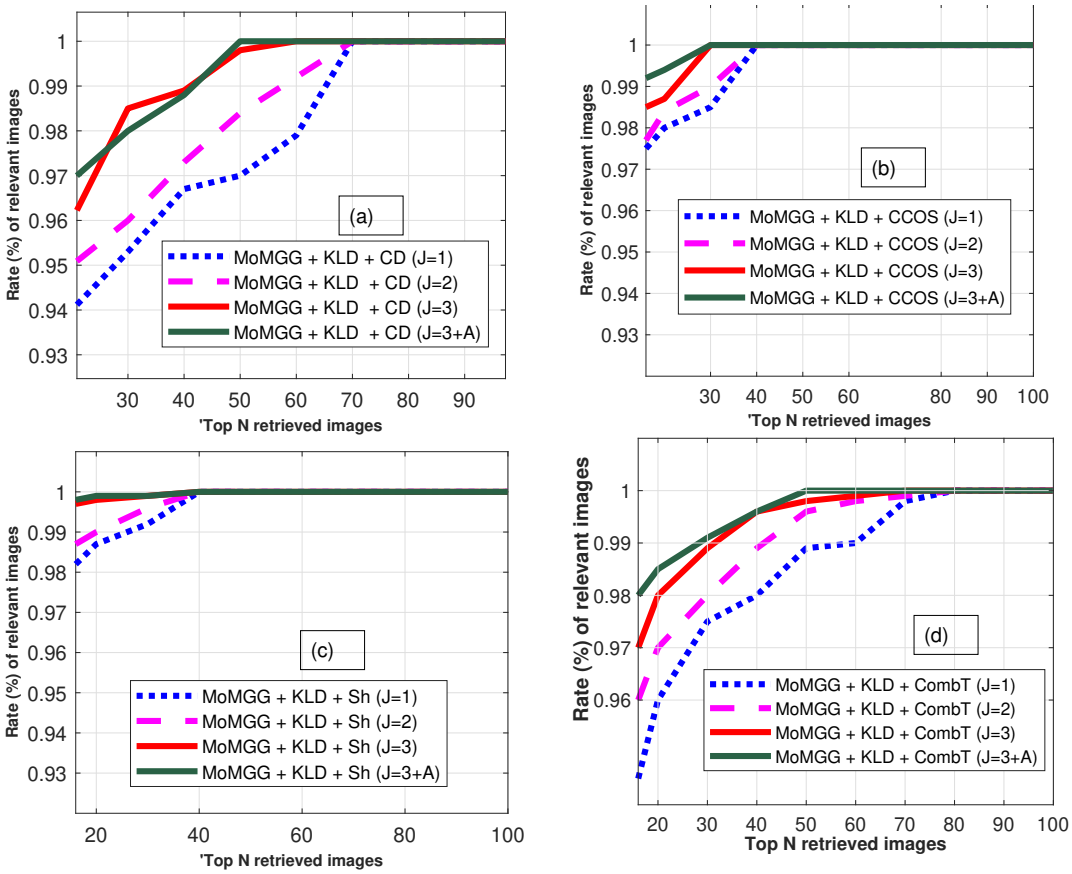


Figure 6.4: Average retrieval rate of 5000 queries as function of the number of top matches considered. (a) Model 1. (b) Model 2. (c) Model 3. (d) Model 4

6.5.2 Comparative evaluation with other methods

To further illustrate the performance of the proposed method, we compared our method with three state-of-the-art methods using the same experimental setting. We used the three benchamark texture datasets are ALOT [218], containing 250 different texture classes with 100 images per class, VisTex(small) containing 40 textures and VisTex (full) which consists of 167 textures. We compute the Average Retrieval Rate (ARR) which is defined as the

average percentage of the retrieved subimages from the same original texture class in the top 16 retrieved subimages. The ARR of the sub-images is given by:

$$ARR = \frac{1}{N_t}\sum_{k=1}^{N_t}\frac{r_{N_r}(I_k)}{N_r}$$

where N_t denotes the total number of sub-images in the dataset, I_k represents the k_th query sub-image, and $r_{N_r}(I_k)$ is the query function which represents the number of correctly retrieved sub-images corresponding to I_k among the N_t retrieved sub-images (i.e. sub-images that belong to a unified image along with the query image I_k). In the experiments, three-scale decomposition is adopted.

Method	ALOT	VisTex(Full)	VisTex(Small)
MoMGG + KLD + CCOS (Scenario 2)	**99,5**	**99,8**	**100**
MoMGG + KLD + CD (Scenario 1)	**96,4**	**97,4**	**99.8**
GC-GG NSST/JD [192]	72,82	69,47	97,02
GC-tLS NSST/JD [192]	70,95	68,43	96,21
MPE/Geodesic [26]	50,1	70,1	91,5
Gabor Wavelet-Copula/KLD [193]	60,8	66,1	92,4

Table 6.2: ARR values of our proposed method along with the existing state-of-the-art methods on the three ALOT, VisTex(Full), and VisTex(Small) datasets. Scenario 1 (accounting the inter-band correlation), Scenario 3 (accounting the spatial distribution)

Table 6.2 compares the ARR that a query retrieved its class images. As can be seen, our proposed MoMGG+KLD+CD (scenario 1) and MoMGG+KLD+CCOS (scenario 2) already obtained the highest ARR values in the three datasets ALOT, VisTex(Full) and VisTex(Small). Moreover, scenario 3 which takes account spatial distribution in texture exhibit impressive retrieval performances. Loosely, taking into account full covariance between coefficients with spatial location, scale and orientation within multi-resolution transform, and multi-modality made a significant difference in the quality of the descriptors, which significantly improves the color texture retrieval.

6.6 Image Reconstruction

This experiment has the objective of evaluating the representativeness of the MoMGG model for texture description. The experiment consists of performing the inverse of the multi-scale decomposition from re-sampled high-frequency bands instead of the original ones. That is to say, after estimating the MoMGG models using the high-frequency bands data, we discard these bands and replace them with re-sampled versions from the calculated MoMGG models and use them jointly with the low-resolution image approximation to reconstruct the original image. We performed this experiment using scenario 3 of bands combination with a two level decomposition. For the sake of comparison, we performed also the same reconstruction using the MoGG [138]), MGG [26] and GG [23] models, respectively.

Fig. 6.5 shows qualitative comparison of reconstruction by the MoMGG compared to other models. We can see that the visual quality of the reconstructed images obtained by our method is generally better than the other models.

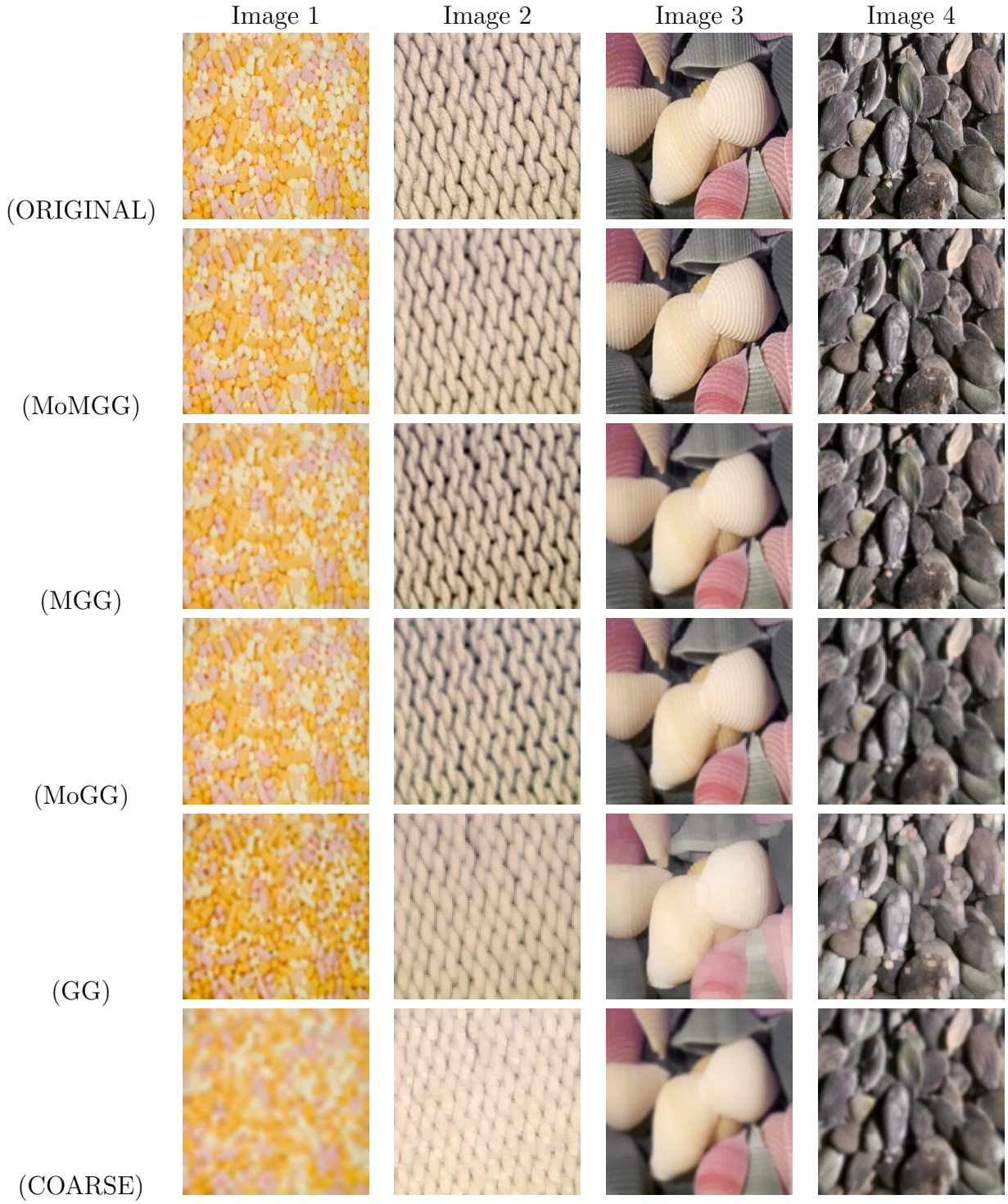


Figure 6.5: Comparison of the visual quality of image reconstruction using sampling from different statistical models. The top row shows the original images, the second to fifth rows show the reconstructed images using the MoMGG (our method), MGG [26], MoGG [138] and GG [23] models, respectively. The last row shows the image approximation used jointly with the re-sampled high-frequency images to reconstruct the original images.

For quantitative evaluation of the reconstruction, we used two popular similarity metrics : the peak-signal-to-noise ratio (PSNR) [232] and the structural similarity index measure (SSIM) [233]. A higher PSNR value provides a higher image quality and a small value of the PSNR implies high differences between images. The SSIM measures the image distortion by combining the loss of luminance, contrast, and correlation between images. The SSIM values range between 0 to 1, where 1 means a perfect match between the reconstructed image and the original one. It is considered to be correlated with the quality perception of the human visual system (HVS). Table 6.3 shows the obtained values of these metrics for the images in Fig. 6.5, while Table 6.4 exhibits their averages for 90 images chosen randomly from the ALOT dataset. We can see that for both metrics, our MoMGG model performed better than the other methods for reconstructing the original image.

	Metric	Image 1	Image 2	Image 3	Image 4
MoMGG (Ours Model)	SSIM	0,982	0,956	0,971	0,940
	PSNR	47,000	45,700	46,220	44,800
MoGG (Allili, 2012)	SSIM	0,850	0,790	0,820	0,780
	PSNR	41,000	40,000	40,000	37,350
MGG (Verdoolaege, 2011)	SSIM	0,820	0,800	0,800	0,780
	PSNR	40,000	41,000	39,000	37,000
GG (Min Do, 2002)	SSIM	0,650	0,560	0,511	0,510
	PSNR	34,000	34,000	35,000	33,000
Coarse image	SSIM	0,570	0,520	0,480	0,480
	PSNR	33,000	32,000	33,000	31,000

Table 6.3: Values of the SSIM and PSNR quality metrics for a reconstructed sample of images from the ALOT dataset.

	Quality metric type	Average (90 images)
MoMGG (Ours Model)	SSIM	0,961
	PSNR	46,900
MoGG (Allili, 2012)	SSIM	0,810
	PSNR	39,200
MGG (Verdoolaege, 2011)	SSIM	0,800
	PSNR	38,250
GG (Min Do, 2002)	SSIM	0.554
	PSNR	34,000
Coarse image	SSIM	0,500
	PSNR	31,000

Table 6.4: Average values for SSIM and PSNR metrics for 90 reconstructed sample of images from the ALOT dataset.

6.7 Conclusion

We have proposed a new statistical model, coined MoMGG, for multi-modal and multivariate color texture image representation. This model generalizes several well-known statistical models for building compact statistical signatures that facilitate image search and classification. In addition of its flexibility to represent the distribution of multi-scale decomposition coefficients, it enables arbitrary combinations of decomposition bands to build a richer and more representative texture signatures. These bands could originate from different color channels or from different image transforms with inter-band correlation. On the other hand, the MoMGG enables accurate fitting of heavy-tailed distributions which makes it more parsimonious. We proposed an EM algorithm based on the via Newton-Raphson method for MoMGG parameter estimation. We then validate the model in the context of content-based texture image retrieval by comparing it to other models in well-known datasets. We observed that our approach outperforms most of statistical-based methods in the literature. Furthermore, we conducted experiments on image reconstruc-

tion re-sampling high-frequency bands from the estimated MoMGG models. Again, our model has shown a very good performance in reconstructing images which gives a huge potential for applications such as super-resolution and image/video compression. Finally, our approach can readily be applied for multivariate representation in dynamic video textures and hyper-spectral images.

Chapter 7

Conclusion and discussion

In this thesis, we have proposed a unified statistical model for texture representation and discrimination based on the formalism of generalized Gaussian distribution (GGD). The new model enables to accurately fit heavy-tailed multi-modal and multivariate data which makes it a ideal tool for representing the distribution of multi-scale transforms in texture images. From this property, we have capable at building high-level applications based on texture discrimination, namely fabric defect detection and content-based texture retrieval. In summary we have made two main contributions in this thesis.

In the first contribution s, using the formalism of univariate mixture of generalized gaussian (MoGG), we have developed an algorithm for defect detection in fabric gray-scale images. This algorithm uses block-wise texture discrimination to detect local fabric defects. The algorithm uses supervised learning to discriminate between defect-free fabrics from defective ones, based on RCT-MoGG signatures. Experiments on several fabric images have shown that our algorithm yields a good detection rates and very low false alarms opposite to state-of-the-art methods. By opposite to several existing approaches for fabric defect detection, which effective in only some types of fabrics and/or defect, our method can deal with not only almost all type of pattern, that is motif-based and non motif-based one, but also all type of defects thanks to the RCT+ MoMG. The method enable both

detection and localization of defects.

In the second contribution building on previous works, we extended the univariate mixtures of GGDs (MoGG), to take into account multidimensional and multimodaldata. We generalized the MoGG model to the mixture of multivariate generalized Gaussians (MoMGG) which allows representing and modelling both multimodality and correlation between transform coefficients. Our method unifies several past approaches for statistical representation of multi-scale transforms. Actually, we have proposed a new statistical model, coined MoMGG, for multi-modal and multivariate color texture image representation. This model generalizes several well-known statistical models for building compact statistical signatures that facilitate image search and classification. In addition of its flexibility to represent the distribution of multi-scale decomposition coefficients, it enables arbitrary combinations of decomposition bands to build a richer and more representative texture signatures. These bands could originate from different color channels or from different image transforms with inter-band correlation. On the other hand, the MoMGG enables accurate fitting of heavy-tailed distributions which makes it more parsimonious. We proposed an EM algorithm based on the via Newton-Raphson method for MoMGG parameter estimation. We then validate the model in the context of content-based texture image retrieval by comparing it to other models in well-known datasets. We observed that our approach outperforms most of statistical-based methods in the literature. Furthermore, we conducted experiments on image reconstruction re-sampling high-frequency bands from the estimated MoMGG models. Again, our model has shown a very good performance in reconstructing images which gives a huge potential for applications such as super-resolution and image/video compression. Finally, our approach can readily be applied for multivariate representation in dynamic video textures and hyper-spectral images.

Bibliography

[1] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, vol. 290, no. 5802, pp. 91-97, 1981.

[2] D. Forsyth, J. Ponce. Computer Vision: A Modern Approach, 2nd Edition, 2012.

[3] Vision Texture Database, *MIT Media Laboratory vision and modeling group.* http://vismod.media.mit.edu/vismod/imagery/VisionTexture

[4] M. Tuceryan, A.K. Jain. Texture analysis. *Handbook Pattern Recognition and Computer Vision*, pp. 235-276, 1993.

[5] A. Depeursinge, O. Al-Kadi, J. Mitchell. Biomedical Texture Analysis. *Academic Press* , 2017.

[6] C. Xie, M. Mirmehdi. TEXEMS: texture exemplars for defect detection on random textured surfaces. *In IEEE Trans Pattern Analysis and Machine Intelligence*, pp. 1454-1464, 2007.

[7] M. S. Allili. Wavelet-Based Texture Retrieval Using a Mixture of Generalized Gaussian Distributions. *20th Int'l Conference on Pattern Recognition*, pp. 3143-3146, 2010.

[8] C. He, S. Li, Z. Liao, M. Liao. Texture classification of PolSAR data based on sparse coding of wavelet polarization textons. *In IEEE Trans Geoscience and Remote Sensing*, pp. 4576-4590, 2013.

[9] L. Gatys, A. Ecker, M. Bethge. Image style transfer using convolutional neural networks. *In CVPR*, pp. 2414-2423, 2016.

[10] R. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5): 786 - 804, 1979.

[11] M. Tuceryan and A. K. Jain. Texture segmentation using Voronoi polygons,. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 211-216, 1990.

[12] A. Rosenfeld and J. Weszka. Picture Recognition. *In Digital Pattern Recognition*, Springer-Verlag, 135-166, 1980.

[13] J. Daugman. Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimised by Two-Dimensional Visual Cortical Filters. *Journal of the Optical Society of America*, 2, 1160-1169, 1985.

[14] A. Bovik, M. Clark and W. Giesler,. Multichannel Texture Analysis Using Localised Spatial Filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12, 55-73, 1990.

[15] S. G. Mallat. Multifrequency channel decompositions of images and wavelet models, Acoustics, Speech and Signal Processing, *IEEE Transactions on*, vol. 37, pp. 2091-2110, 1989.

[16] J. S. De Bonet. Multiresolution sampling procedure for analysis and synthesis of texture images. *In Proceedings of the 24th annual conference on Computer graphics and interactive techniques, ACM SIGGRAPH Comput Graph*, 1997.

[17] M. N. Do and M. Vetterli. Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models. *In IEEE Transactions on Multimedia*, pp. 517-527, 2002.

[18] S.K. Choy, C.S. Tong and J. Math. Supervised Texture Classification Using Characteristic Generalized Gaussian Density. *Journal of Mathematical Imaging and Vision*, pp 35-47, 2007.

[19] R. M. Haralick, SHANMUGAM K., DINSTEIN I., Textural features for image classification. *IEEE Transactions on Systems, Man et Cybernetics, SMC-3*, pp. 610-621, 1973.

[20] D. A. Al Nadi, A. M. Mansour. Independent Component Analysis (ICA) for Texture Classicication. *IEEE int'l Multi-conf. on Systems, Signals, and devices*. pp. 1-5, 2008.

[21] Wainwright, J. Martin and E. P. Simoncelli. Scale mixtures of Gaussians and the statistics of natural images. *In Advances in neural information processing systems*, pp. 855-861. 2000

[22] J. Portilla and E. P. Simoncelli. Parametric texture model based on wavelet coeficients. *International Journal of Computer Vision*, pp. 49-71, 2000.

[23] M. Do and M. Vetterli , Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance, *IEEE Trans. Image Processing*, pp. 146 - 158, 2002.

[24] G. Van De Wouwer, P. Scheunders, D. Van Dyck Statistical Texture Characterization From Discrete Wavelet Representations. *IEEE Transactions on Image Processing*, 8(4) : 592-598, 1999.

[25] G. Verdoolaege, Y. Rosseel, M. Lambrechts and P. Scheunders. Wavelet-based colour texture retrieval using kullback-Leibler divergence between bivariate generalized gaussian models. *In Proceedings of the IEEE international conference on image processing*, pp. 265 - 268. Cairo, 2009.

[26] G. Verdoolaege and P. Scheunders. Geodesics on the manifold of multivariate generalized gaussian distributions with an application to multicomponent texture discrimination. *International Journal of Computer Vision*, 95(3):265-276, 2011.

[27] G. Verdoolaege and P. Scheunders. On the geometry of multivariate generalized gaussian models. *Journal of Mathematical Imaging and Vision*, pp 180-193, 2012.

[28] L. Bombrun, Y. Berthoumieu, N. Lasmar, and G. Verdoolaege. Multivariate texture retrieval using the geodesic distance between elliptically distributed random variables. *In 18th IEEE International Conference on Image Processing (ICIP)*, pp. pp. 3637-3640, 2011.

[29] Kwitt, R. et Uhl, A., A joint model of complex wavelet coefficients for texture retrieval. *In Proceedings of IEEE International Conference on Image Processing - ICIP*, pp. 1877-1880, 2009.

[30] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa and M. Pietikäinen. From BoW to CNN: Two Decades of Texture Representation for Texture Classification. *Int'l J. of Computer Vision*, 127(1):74-109, 2018.

[31] L. Gatys, A. Ecker, M. Bethge. Image style transfer using convolutional neural networks. *IEEE Conf. on Computer Vision and Pattern Recognition*, 2414-2423, 2016.

[32] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *Int'l Conf. on Learning Representations*, 1-14, 2015.

[33] T.S.A. Wallis, C.M. Funke, A.S. Ecker, L.A. Gatys, F.A. Wichmann and M. Bethge. A Parametric Texture Model Based on Deep Convolutional Features Closely Matches Texture Appearance for Humans. *Journal of Vision*, 17(12):5, 1-29, 2017

[34] Brodatz Texture Images. http ://www.ux.uis.no/tranden/brodatz.html

[35] R. Kwitt, P. Meerwald and A. Uhl. Efficient Texture Image Retrieval Using Copulas in a Bayesian Framework. *IEEE Trans. on Image Processing*, 20(7):2063-2077, 2011.

[36] R. Jenssen, J.C. Principe, T. Eltoft. Cauchy–Schwartz pdf divergence measure for non-parametric clustering. *IN Proceedings of IEEE Norway Section Signal Processing Symposium*, pp. 2–4, 2003.

[37] R. Jenssen, D. Erdogmus, K.E. Hild, J.C. Principe, T. Eltoft. Optimizing the Cauchy–Schwarz PDF distance for information theoretic, non-parametric clustering. *In Proceedings of International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 34–45, 2005.

[38] B. Julesz. Visual Pattern Discrimination. *In IRE Transactions on Information Theory*, pp. 84-92, 1962.

[39] A. Gagalowicz. A new method for texture fields synthesis : Some applications to the study of human vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(5): pp. 520-533, 1981.

[40] G. R. Cross, A. K. Jain. Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 25-39, 1983.

[41] F. W Campbell, J. G. Robson. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, pp. 551-566, 1968.

[42] R. L. De Valois, D. G. Albrecht, L. G. Thorell, Spatial-frequency selectivity of cells in macaque visual cortex. *Vision Research*, pp. 545-559, 1982.

[43] T.B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer vision and image understanding*, pp. 231-268, 2001.

[44] H. Donald and M. Pauline Baker (1994). Computer Graphics. *Prentice-Hall*, 1994.

[45] S.Davis. Computer graphics. New York, 2011.

[46] M. Tuceryan, A. K. Jain. Texture analysis. *The handbook of pattern recognition and computer vision*, pp. 207-248, 1998.

[47] G. Eichmann and T. Kasparis. Topologically invariant texture descriptors. *Computer Vision Graphics and Image Processing*, 41(3), pp. 267-281, 1988.

[48] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), pp. 51-59, 1996.

[49] T. Ojala, M. Pietikäinen and T. Mäenpää. A Generalized Local Binary Pattern Operator for Multiresolution Gray Scale and Rotation Invariant Texture Classification. *In Advances in Pattern Recognition*, pp. 399-408, 2001.

[50] T. Ojala, M. Pietikäinen, et T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 971-987, 2002.

[51] M. Pietikainen, A. Hadid, G. Zhao, and Y. Ahonen. Computer Vision Using Local Binary Patterns. *Computational Imaging and Vision*. Springer-Verlag, 2011.

[52] E. G. Simona, N. Petkov, and P. Kruizinga. Comparison of texture features based on gabor filters. *IEEE Transactions on image processing*, 11(10):1160-1167, 2002.

[53] F. L. Hang and M. D. Levine. Finding a small number of regions in an image using low-level features. *Pattern Recognition*, pp. 2323-2339, 2001.

[54] Y. Rubner, J. Puzicha, C.Tomasi and J. M. Buhmann. Empirical evaluation of dissimilarity measures for color and texture. *Computer vision and image Understanding*, pp. 25-43, 2001.

[55] Z. Zhang. Feature-based facial expression recognition : Experiments with a multi-layer perception. *Rapport de Recherche INRIA*, 1998.

[56] G. Tzagkarakis, B. Beferull-Lozano, and P. Tsakalides. Rotation-invariant texture retrieval with gaus- sianized steerable pyramids. *IEEE Transactions on Image Processing*, 15(9), pp. 2702-2718, 2006.

[57] J. Morlet, G. Arens, E. Fourgeau, and D. Glard. Wave propagation and sampling theory—Part I: Complex signal and scattering in multilayered media. *In Geophysics*, pp. 203-221, 1982.

[58] M. Unser. Texture classification and segmentation using wavelet frames. *In IEEE Transactions on Image Processing*, pp. 1549-1560, Nov. 1995.

[59] A.F. Laine and J. Fan. Frame representations for texture segmentation. *IEEE transactions on image processing*, 5(5), 771-780, 1996.

[60] S.C. Kim and T.J. Kang. Texture classification and segmentation using wavelet packet frame and Gaussian mixture model. *Pattern Recognition*, 40(4), pp. 1207-1221, 2007.

[61] Y.-L. Qiao, Z.-M. Lu, J.-S. Pan et S.-H. Sun. Fast k-nearest neighbor search algorithm based on pyramid structure of wavelet transform and its application to texture classification. *Digital Signal Processing*, pp. 837-845, 2010.

[62] S. Gai, G. Yang and S. Zhang. Multiscale texture classification using reduced quaternion wavelet transform. *AEU - International Journal of Electronics and Communications*, pp. 233-241, 2013.

[63] P. De Rivaz and N. Kingsbury. Complex wavelet features for fast texture image retrieval. *In 1999 International Conference on Image Processing*, pp. 109-113, 1999.

[64] R. Kwitt and A. Uhl. Lightweight Probabilistic Texture Retrieval. I*n IEEE Transactions on Image Processing*, pp. 241-253, 2010.

[65] Selensky I. W., Baraniuk R. G., Kingsbury N. C., The dual-tree complex wavelet transform, *IEEE Signal Processing Magazine*, pp. 123-151, 2005.

[66] G. Peyre. Géométrie multi-échelles pour les images et les textures. *Thèse de doctorat, Ecole Polytechnique X*, 2005.

[67] D. D. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. *In IEEE Transactions on Image Processing*, pp. 1610-1620, 2006.

[68] S. Arivazhagan, L. Ganesan and T. G. Subash Kumar. Texture classification using ridgelet transform. *Pattern Recognition Letters*, 27(16), pp. 875-1883, 2006.

[69] L. Dettori and L. Semler. A comparison of wavelet, ridgelet, and curvelet-based texture classification algorithms in computed tomography. *Computers in Biology and Medicine*, 37(4) , pp. 486-498, 2007.

[70] M. N Do, and M. Vetterli. Contourlets: A Directional Multiresolution Image Representation. I*EEE Int'l Conf. on Image Processing*, pp. 357-360, 2002.

[71] M. N Do, and M. Vetterli. The Contourlet Transform : an Efficient Directional Multiresolution Image Representation. *IEEE Trans. on Image Processing*, pp. 2091- 2106, 2005.

[72] B. Julesz, E. N. Gilbert, J. D. Victor. Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics,*. pp. 137-140, 1978.

[73] M. Sonka, V. Hlavac and R. Boyle. Image Processing. I*n Chapman & Hall on Analysis and Machine Vision*, 1993.

[74] S.Herlidou. Caractérisation tissulaire en Imagerie par Résonnance Magnétique Nucléaire par l'analyse de texture : étude du tissu musculaire et de tumeurs intracraniennes. *PhD thesis, Université de Rennes 1*, 1999.

[75] M. Sharma, S. Singh. Evaluation of texture methods for image analysis. *Department of Computer Science , University of Exeter*, 2000.

[76] . M. L. Comer and E. J. Delp. Segmentation of textured images using a multiresolution Gaussian autoregressive model. *in IEEE Transactions on Image Processing*, pp. 408-420, 1999.

[77] O. Alata and C. Ramananjarasoa. Unsupervised textured image segmentation using 2-d quarter plane autoregressive model with four prediction supports. *Pattern RecognitionLetters*, 26(8):1069-1081, 2005.

[78] I. U. H. Qazi, I.-U.-H., O. Alata, J.-C. Burie, M. Abadi, A. Moussa and C. Fernandez-Maloigne. Parametric models of linear prediction error distribution for color texture and satellite image segmentation. *Computer Vision and Image Understanding*, 115(8), pp. 1245-1262, 2011.

[79] M. D. Jolly and A. Gupta. Color and texture fusion: application to aerial image segmentation and GIS updating. *Proceedings Third IEEE Workshop on Applications of Computer Vision*, pp. 2-7, 1996.

[80] C. Bouman and B. Liu. Multiple resolution segmentation of textured images. I*EEE Transactions on pattern analysis and machine intelligence*, 39(9), pp. 2025-2043, 1991.

[81] F. Cohen. Classification of rotated and scaled texture images using Gaussian Markov random field models. *IEEE Trans.* 13(2), pp. 192-202, 1992.

[82] J. L. Chen and A. Kundu. Rotation and gray scale transform invariant texture recognition using hidden Markov model. *Proceedings of the ICASSP*, pp.69-72, 1992.

[83] J.-L. Chen and A. Kundu. Rotation and grey-scale transform invariant texture identification using wavelet decomposition and HMM. *In IEEE Trans.* 16(2), pp. 208-214, 1994.

[84] J.-L. Chen and A. Kundu. Unsupervised texture segmentation using multichannel decomposition and hidden Markov models. *In IEEE Trans. Image Process.* 4 (5), pp. 603-619, 1995.

[85] M. Pi, C. S. Tong, S. K. Choy, and H. Zhang. A fast and effective model for wavelet subband histograms and its application in texture image retrieval. *In IEEE Trans. Image Process.*, pp.3078-3088, 2006.

[86] S. K. Choy and C. S. Tong, Statistical properties of bit-plane probability model and its application in supervised texture classification. *In IEEE Trans. Image Process.*, pp. 1399-1405, 2008.

[87] S. K. Choy and C. S. Tong. Statistical Wavelet Subband Characterization Based on Generalized Gamma Density and Its Application in Texture Retrieval. *In IEEE Transactions on Image Processing*, pp. 281-289, 2010

[88] A. Suzuki and H. Shouno. A Generative Model of Textures Using Hierarchical Probabilistic Principal Component Analysis. *arXiv preprint arXiv:1810.06892*, 2018.

[89] M. S. Allili, N. Bouguilla and D. Ziou. Finite General Gaussian Mixture Modelling and Application to Image and Video Forground segmentation. *Journal of Electronic imaging*, 17 :013005, 2008.

[90] H. Rami, L. Belmerhnia, A. Drissi El Maliani and M. El Hassouni. Texture Retrieval Using Mixtures of Generalized Gaussian Distribution and Cauchy-Schwarz Divergence in Wavelet Domain. *Signal Processing: Image Communication*, pp. 45-58, 2016.

[91] P. Burt and A. Adelso., The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, pp. 532-540, 1983.

[92] M. Turner. Texture discrimination by Gabor functions. *Biological Cybernetics*, pp. 71-82, 1986.

[93] M. Clark and A. Bovik, Texture segmentation using Gabor modulation/ demodulation, Pattern Recognition Letters. 6(4), 261-267, 1987.

[94] A. Kumar and G. Pang, Defect detection in textured materials using Gabor filters. *IEEE Transactions on Industry Applications*, 38(2), 425-440, (2002).

[95] A.C. Bovik, Analysis of multichannel narrow-band 0lters for image texture segmentation, IEEE Trans. Signal Process. 39 (1990) 2025-2043

[96] A.C. Bovik, N. Gopal, T. Emmoth, A. Restrepo. Localised measurement of emergent image frequencies by Gabor wavelets. *IEEE Trans. Inform. Theory*, pp. 691-712, 1992.

[97] T. Randen and J. H. Husoy. Filtering for texture classification: a comparative study. *In IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4), 291-310, 1999.

[98] I. Daubechies. Ten lectures on wavelet. *Society for Industrial and Applied Mathematics*, 1992.

[99] T. Chang and C. J. Kuo. Texture analysis and classification with tree-structured wavelet transform. *In IEEE Transactions on Image Processing*, pp. 429-441, 1993.

[100] R. Porter, N. Canagarajah. Robust rotation invariant texture classiffication. *International Conference on Acoustics, Speech and Signal Processing*, pp. 3157-3160, 1997.

[101] R. Porter, N. Canagarajah, Robust rotation-invariant texture classiffication: wavelet, Gabor filter and GMRF based schemes. *In IEEE Proc. Vision Image Signal Process.* 144(3), pp. 180-188, 1997.

[102] R. Porter and N. Canagarajah. A Robust Automatic Clustering Scheme for Image Segmentation Using Wavelets. *In IEEE Trans. Image Processing*, pp. 662-665, 1996.

[103] P. Burt and A. Adelso., The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, pp. 532-540, 1983.

[104] R. H. Bamberger and M. J. T. Smith. A filter bank for the directional decomposition of images: theory and design. *In IEEE Transactions on Signal Processing*, pp. 882-893, 1992.

[105] J. M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients IEEE Trans. Signal Proc., 41 12 3445 3462 doi: 10.1109/78.258085, 1993.

[106] M. Vetterli. Multidimensional subband coding: Some theory and algorithms. I*n Signal Proc.*, pp. 97-112, 1984.

[107] D. D. -. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. I*n IEEE Transactions on Image Processing*, pp. 1610-1620, 2006.

[108] N. Baaziz, Adaptive watermarking schemes based on a redundant contourlet transform," in Proc. IEEE Int. Conf. Image Process., Sep. 2005, pp. 221–224.

[109] M.S. Allili, N. Baaziz and M. Mejri. Texture Modeling Using Contourlets and Finite Mixtures of Generalized Gaussian Distributions and Applications. *IEEE Trans. on Multimedia*, 16(3):772-784, 2014.

[110] B. Aiazzi, L. Alparone, and S. Baronti. Estimation based on entropy matching for generalized Gaussian PDF modelling. *IEEE Signal Process.* pp. 138-140, 1999.

[111] G.-E. P. Box et G. C. Tiao. *Bayesian Inference in Statistical Analysis*. Wiley, 1992.

[112] P. Moulin and J. Liu. Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity schemes. *IEEE Trans. Inf. Theory*, pp. 909-919, 1999.

[113] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Processing*, pp. 1532-1546, 2000.

[114] K. Sharifi and A. Leon-Garcia. Estimation of shape parameter for generalized Gaussian distribution in subband decomposition of video. *IEEE Trans. Circuits Syst. Video Technol.*, pp. 52-56, 1995.

[115] S. Kasaei, M. Deriche, and B. Boashash. A novel fingerprint image compression technique using wavelets packets and pyramid latticevector quantization. *IEEE Trans. Image Process.*, pp. 1365-1378, 2002.

[116] M. Baccar, L.-A. Gee, and M. A. Abidi. Reliable location and regression estimates with application to range image segmentation. *Journal of Mathematical Imaging and Vision*, pp. 195-205, 1999.

[117] S.-K. Choy and C.-S. Tong. Supervised texture classification using characteristic generalized Gaussian density. *Journal of Mathematical Imaging and Vision*, pp. 35-47, 2007.

[118] S. Gazor and Wei Zhang. Speech probability distribution. *In IEEE Signal Processing Letters*, 10(7), pp. 204-207, 2003.

[119] S. Choi, A. Cichocki, S. Amari. Flexible independent component analysis. *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, 26 (1) 25, 2000.

[120] K. Kokkinakis, A.K. Nandi. K. Kokkinakis and A. K. Nandi. Optimal blind separation of convolutive audio mixtures without temporal constraints. *IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Québec*, pp. I-217, 2004.

[121] S. YU, A. ZHANG and H. Li. A review of estimating the shape parameter of generalized Gaussian distribution. *Journal of Computational Information Systems*, pp. 9055-9064, 2012.

[122] E. de Ves, X. Benavent, A. Ruedin, D. Acevedo, and L. Seijas. Wavelet-based texture retrieval and modelling the magnitudes detail coefficients with generalized gamma distribution. *In Proc. IEEE Int. Conf. Pattern Recog*, pp. 221-224, 2010.

[123] J. J. Oliver, R. A. Baxter et C. S. Wallace. Unsupervised Learning Using MML. *Proc. Int. Conf. Mach. Learn.* pp. 364-372, 1996.

[124] C. S. Wallace. Statistical and Inductive Inference by Minimum Message Length. *Information Science and Statistics.*. Springer Verlag, 2005.

[125] M. E. Tipping and C. Bishop. Probabilistic Principal Component Analysis. J*ournal of the Royal Statistical Society*, pp. 611-622, 1999.

[126] I. T. Jolliffe. Principal Component Analysis. New York - Springer, 1986.

[127] F. Pascal, L. Bombrun, J. Y. Tourneret andY. Berthoumieu. Parameter Estimation For Multivariate Generalized Gaussian Distributions. *IEEE trans. on signal processing*, 61(23), 2013.

[128] G. Agro. Maximum likelihood estimation for the exponential power function parameters. *Commun. Statist.-Simul. Computat.*, 24(2), pp. 523- 536, 1995.

[129] M. K. Varanasi and B. Aazhang. Parametric generalized Gaussian distributiontion. *J. Acoust. Soc. Amer.*, 86(4), pp.404-415, 1989.

[130] C.S. Cho and B.M. Chung and M.J. Park. Development of Real-Time Vision-Based Fabric Inspection System. *IEEE Trans. on Industrial Electronics*, 52(4):1073-1079, 2005.

[131] N. Khelil-Cherif and A. Benazza-Benyahia. Wavelet-based multivariate approach for multispectral image indexing. *In SPIE Conf. Wavelet Appl. Indust. Process.*, Philadelphia, PA, USA, 2004.

[132] T. Zhang, A. Wiesel, and M. Greco. Multivariate generalized Gaussian distribution: Convexity and graphical models. *IEEE Trans. Signal Process.*, pp. 414 - 4148, 2013.

[133] F. Pascal, Y. Chitour, J. P. Ovarlez, P. Forster, and P. Larzabal. Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis. *IEEE Trans. Signal Process.*, pp. 34-48, 2008.

[134] Y. Chitour and F. Pascal. Exact maximum likelihood estimates for SIRV covariance matrix: Existence and algorithm analysis. *IEEE Trans. Signal Process.*, pp. 4563-4573, 2008.

[135] R. A. Maronna. Robust M-estimators of multivariate location and scatte. *Ann. Statist.*, pp. 51-67, 1976.

[136] R. Cossu, I. H. Jermyn, and J. Zerubia. Texture analysis using probabilistic models of the unimodal and multimodal statistics of adaptive wavelet packet coefficients. *In Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 597-600, 2004.

[137] S. Kullback. Information theory and statistics. Newyork:Docer, 1968.

[138] M. S. Allili. Wavelet Modeling Using Finite Mixtures of Generalized Gaussian Distributions: Application to Texture Discrimination and Retrieval. *EEE Transactions on Image Processing*, pp.1452,1464, 2012.

[139] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Texts in Statistics, 2004.

[140] Workgroup on Texture Analysis of DFG (http://lmb.informatik.uni-freiburg.de/research/dfg-texture/tilda).

[141] A. P. Dempster, N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*. pp. 1–38, 1977.

[142] A. Kumar. Computer-Vision-Based Fabric Defect Detection: A Survey. *IEEE Trans. on Industrial Electronics*, 55(1):348-363, 2008.

[143] H.Y.T. Ngan, G.K.H. Pang and N.H.C. Yung. Automated Fabric Defect Detection: A Review. *Image and Vision Computing*, 29(7):442-458, 2011.

[144] D. Yapi, M. Mejri and M.S. Allili and N. Baaziz. A Learning-Based Approach for Automatic Defect Detection in Textile Images. *IFAC Symposium on Information Control in Manufacturing*, 2423-2428, 2015.

[145] D.E. Joyce. Walpaper groups: (*http://www2.clarku.edu/ djoyce/wallpaper/seventeen.html*), 1994.

[146] H.Y.T. Ngan, G.K.H. Pang and N.H.C. Yung. Performance Evaluation for Motif-Based Patterned Texture Defect Detection. *IEEE Trans. on Automation Science and Engineering*, 7(1):58-72, 2010.

[147] C-F.J. Kuo and T-L. Su. Gray Relational Analysis for Recognizing Fabric Defects. *Textile Research Journal*, 73(5):461-465, 2003.

[148] M.K. Ng, H.Y.T. Ngan, X. Yuan and W. Zhang. Patterned Fabric Inspection and Visualization by the Method of Image Decomposition. *IEEE Trans. on Automation Science and Engineering*, 11(3):943-947, 2014.

[149] H.Y.T. Ngan, G.K.H. Pang, S.P. Yung and M.K. Ng. Wavelet-Based Methods on Patterned Fabric Defect Detection. *Pattern Recognition*, 38(4):559-576, 2005.

[150] H.Y.T. Ngan and G.K.H. Pang. Regularity Analysis for Patterned Texture Inspection. *IEEE Trans. on Automation Science and Engineering*, 6(1):131-144, 2009.

[151] G. Stübl, J-L. Bouchot, Pe. Haslinger and B. Moser. Discrepancy Norm as Fitness Function for Defect Detection on Regularly Textured Surfaces. *DAGM/OAGM Symposium on Pattern Recognition*, LNCS 7476, pp. 428–437, 2012.

[152] C.S.C. Tsang, H.Y.T. Ngan and G.K.H.Pang. Fabric inspection Based on the Elo rating Method. *Pattern Recognition*, 51(17):378–394, 2016.

[153] H.Y.T. Ngan, G.K.H. Pang, and N.H.C. Yung. Motif-Based Defect Detection for Patterned Fabric. *Pattern Recognition*, 41(6):1878-1894, 2008.

[154] N. Baaziz. Adaptive Watermarking Schemes Based on a Redundant Contourlet Transform. *IEEE Int'l Conf. on Image Processing*, 221-224, 2005.

[155] M. Bennamoun and A. Bodnarova. Automatic Visual Inspection and Flaw Detection in Textile Materials: Past, Present and Future. *IEEE Int'l Conf. on Systems, Man, and Cybernetics*, 4340-4343, 1998.

[156] R.T. Chin and C.A. Harlow. Automated Visual Inspection: A Survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 4(6):557-573, 1982.

[157] A.L. Amet and A. Ertüzün and A. Erçi. An Efficient Method for Texture Defect Detection: Sub-band Domain Co-occurrence Matrices. *Image and Vision Computing*, 18(6-7):543-553, 2000.

[158] C.H. Chan and G. K.H. Pang. Fabric Defect Detection by Fourier Analysis. *IEEE Trans. on Industry Applications*, 36(5):1267-1276, 2000.

[159] M.S. Allili and N. Baaziz: Contourlet-Based Texture Retrieval Using a Mixture of Generalized Gaussian Distributions. *Computer Analysis of Images and Patterns*, 446-454, 2011.

[160] A. Boulmerka and M.S. Allili. Thresholding-Based Segmentation Revisited Using Mixtures of Generalized Gaussian Distributions. *IEEE Int'l Conf. on Pattern Recongition*, 2894-2897, 2012.

[161] F.S. Cohen, Z. Fan and S. Attali. Automated Inspection of Textile Fabrics Using Textural Models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(8):803-808, 1991.

[162] M. N. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Trans. on Image Processing*, 14(12):2091-2106, 2005.

[163] A.E. Elo. *The Rating of Chess Players, Past and Present*. Ishi Press, 2008.

[164] S. Ghorai, A. Mukherjee, M. Gangadaran and P.K. Dutta. Automatic Defect Detection on Hot-Rolled Flat Steel Products. *IEEE Trans. on Instrumentation and Measurement*, 62(3):612-621, 2013

[165] J-F. Jing, S. Chen and P-F. Li. Fabric Defect Detection Based on Golden Image Subtraction. *Coloration Technology*, 132:1-14, 2016.

[166] A. Kumar and G.K.H. Pang. Defect Detection in Textured Materials Using Optimized Filters. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 32(5):553-570, 2002.

[167] A. Kumar. Neural Network Based Detection of Local Textile Defects. *Pattern Recognition*, 36(7):1645-1659, 2003.

[168] C-F.J. Kuo, C-J. Lee and C-C. Tsai. Using a Neural Network to Identify Fabric Defects in Dynamic Cloth Inspection. *Textile Research Journal*, 73(3):238-244, 2003.

[169] C-F.J. Kuo and C.J. Lee. A Back-Propagation Neural Network for Recognizing Fabric Defects. *Textile Research Journal*, 73(2):147-151, 2003.

[170] R.A. Lizarraga-Morales, R.E. Sanchez-Yanez and R. Baeza-Serrato. Defect Detection on Patterned Fabrics Using Texture Periodicity and the Coordinated Clusters Representation. *Textile Research Journal*, DOI: 10.1177/0040517516660885, 1-14, 2016.

[171] H. Lee, A. Battle, R. Raina and A.Y. Ng. Efficient Sparse Coding Algorithms. *Advances in Neural Information Processing Systems*, 801-808, 2006.

[172] M. Li, Z-M. Deng and L. Wang. Defect Detection of Patterned Fabric by Spectral Estimation Technique and Rough Set Classifier. *Global Congress on Intelligent Systems (GCIS)*, 190-194, 2013.

[173] A.F. Limas Serafim. Segmentation of Natural Images Based on Multiresolution Pyramids Linking of the Parameters of an Autoregressive Rotation Invariant Model: Application to Leather Defects Detection. *IAPR/IEEE Int'l Conf. on Pattern Recognition*, 41–44, 1992.

[174] G. McLachlan and D. Peel. *Finite Mixture Models*. John Wiley & Sons, 2005.

[175] M. Mirmehdi, X. Xie and J. Suri. *Handbook of Texture Analysis*. World Scientific, 2008.

[176] S. Ozdemir, A. Ercil. Markov Random Fields and Karhunen–Loeve Transform for Defect Inspection of Textile Products. *IEEE Conf. Emerging Technologies & Factory Automation*, 697–703, 1996.

[177] G.K.H. Pang. Novel Method for Patterned Fabric Inspection Using Bollinger Bands, *Optical Engineering*, DOI:10.1117/1.2345189, 2006.

[178] S.J. Raudys and A.K. Jain. Small Sample Size Effects in Statistical Pattern Recognition: Recommendations for Practitioners. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(3):252-264, 1991.

[179] H. Sari-Sarraf and J.S. Goddard. Vision System for On-loom Fabric Inspection. *IEEE Trans. on Industry Applications*, 35(6):1252-1259, 1999.

[180] O.G. Sezer and A. Ertuzun and A. Ercil. Using Perceptual Relation of Regularity and Anisotropy in the Texture with Independent Component Model for Defect. *Pattern Recognition*, 40(1):121-133, 2007.

[181] F. Tajeripour, E. Kabir and A. Sheikhi. Fabric Defect Detection Using Modified Local Binary Patterns. *EURASIP J. on Advances in Signal Processing*, ID 783898, 2008.

[182] I-S. Tsai, C.H. Lin and J-J. Lin. Applying an Artificial Neural Network to Pattern Recognition in Fabric Defects. *Textile Research Journal*, 65(3):123-130, 1995.

[183] D-M. Tsai and T-Y. Huang. Automated Surface Inspection for Statistical Textures. *Image and Vision Computing*, 21(4):307-323, 2003.

[184] Z. Wang. A New Approach for Segmentation and Quantification of Cells or Nanoparticles. *IEEE Trans. on Industrial Informatics*, 12(3):962-971, 2016.

[185] Z. Wang. A Semi-Automatic Method for Robust and Efficient Identification of Neighboring Muscle Cells. *Pattern Recognition*, 53:300-312, 2016.

[186] E.J. Wood. Applying Fourier and Associated Transforms to Pattern Characterization in Textiles. *Textile Research Journal*, 67(4):2012-2020, 1990.

[187] X. Yang, K. Pang and N. Yung. Discriminative Fabric Defect Detection Using Adaptive Wavelets. *Optical Engineering*, 41(12):3116-3126, 2002.

[188] X. Yang and G. Pang and N. Yung. Discriminative Training Approaches to Fabric Defect Classification Based on Wavelet Transform. *Pattern Recognition*, 37(5):889-899, 2004.

[189] J. Zhou, D. Semenovich, A. Sowmya and J. Wang. Sparse Dictionary Reconstruction for Textile Defect Detection. *IEEE Int'l Conf. on Machine Learning and Applications*, 21-26, 2012.

[190] J. Zhou and J. Wang. Fabric Defect Detection Using Adaptive Dictionaries. *Textile Research Journal*, 83(17):1846-1859, 2013.

[191] Q. Zhu, M. Wu, J. Li and D. Deng. Fabric Defect Detection via Small Scale Over-complete Basis Set. *Textile Research Journal*, 84(15): 1634-1649, 2014.

[192] S. Etemad, M. Amirmazlaghani,. Color Texture Image Retrieval Based on Copula Multivariate Modeling in the Shearlet Domain. arXiv preprint arXiv:2008.00910, 2020.

[193] C. Li , Y. Huang , L. Zhu , Color texture image retrieval based on gaussian copula models of gabor wavelets, Pattern Recognit. 64, 118–129, 2017.

[194] A. Humeau-Heurtier. Texture feature extraction methods: A survey. IEEE Access, 7, 8975-9000, 2019.

[195] E. Bazan, P. Dokladal, E. Dokladalova. Quantitative Analysis of Similarity Measures of Distributions. *British Machine Vision Conference*, 1-12, 2019.

[196] E. Simoncelli, J. Portilla, Texture characterization via joint statistics of wavelet coefficient magnitudes. In Fifth IEEE Int'l Conf. on Image Proc., Chicago, October 4–7, Vol. I. IEEE Computer Society, 1998.

[197] M. Mehri, P. Gomez-Krämer, P. Héroux, M. Coustaty, J. Lerouge and R. Mullot. A bottom-up method using texture features and a graph-based representation for lettrine recognition and classification, 13th International Conference on Document Analysis and Recognition (ICDAR), Tunis, 226-230, 2015.

[198] A. Mehnert, R. Moshavegh, K. Sujathan, P. Malm, and E. Bengtsson, A structural texture approach for characterising malignancy associated changes in pap smears

based on mean-shift and the watershed transform, in Proc. 22nd Int. Conf. Pattern Recognit., Art. no. 6976924, 2014.

[199] Y. Dong, D. Tao, X. Li, J. Ma, and J. Pu, Texture classification and retrieval using shearlets and linear regression," IEEE Trans. Cybern., vol. 45, no. 3, pp. 358-369, Mar. 2015.

[200] J. Chen et al., Robust local features for remote face recognition, Image Vis. Comput., vol. 64, pp. 34-46, Aug. 2017.

[201] T. Song, H. Li, F. Song, Q.Wu, and J. Cai, LETRIST: Locally encoded transform feature histogram for rotation-invariant texture classification, IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 7, pp. 1565-1579, Jul. 2018.

[202] J. Soldera, C. T. J. Dodson, J. Scharcanski, Face recognition based on texture information and geodesic distance approximations between multivariate normal distributions, Meas. Sci. Technol., vol. 29, no. 11, p. 114001, 2018.

[203] A. Ramola, A. K. Shakya, D. Van Pham.. Study of statistical methods for texture analysis and their modern evolutions. Engineering Reports, 2(4), e12149, 2020.

[204] Yu Song, Changsheng Chen, MPPCANet: A feedforward learning strategy for few-shot image classification, Pattern Recognition, vol.113, p. 107792, 2021.

[205] E.P Simoncelli and B.A Olshausen. Natural Image Statistics and Neural Reprsentation. Annual Review of Neuroscience, 24:1193–1216, 2001.

[206] M. Kivanc Mihcak;I. Kozintsev;K. Ramchandran;P. Moulin Low-complexity image denoising based on statistical modeling of wavelet coefficients. IEEE Signal Process Lett 6(12):300–303, 1999.

[207] A. Karine, A.D.E. Maliani, M.E. Hassouni, A novel statistical model for content-based stereo image retrieval in the complex wavelet domain, Journal of Visual Communication and Image Representation, Volume 50, 2018, Pages 27-39.

[208] D. Yapi, M. S. Allili and N. Baaziz, "Automatic Fabric Defect Detection Using Learning-Based Local Textural Distributions in the Contourlet Domain," in IEEE Transactions on Automation Science and Engineering, vol. 15, no. 3, pp. 1014-1026, 2018.

[209] G. Huang, Z. Liu, K.Q. Weinberger, L. Van der Maaten. Densely connected convolutional networks. *IEEE Computer Vision and Pattern Recognition*, 4700-4708, 2017.

[210] J. Freeman, C. Ziemba, D.J. Heeger, E.P. Simoncelli and J.A. Movshon. A Functional and Perceptual Signature of the Second Visual Area in Primates. *Nature Neuroscience*, 16(7):974-981, 2013.

[211] G. Okazawa, S. Tajima, H. Komatsu. Image Statistics Underlying Natural Texture Selectivity of Neurons in Macaque V4. *Proc. of the National Academy of Sciences*, 112(4):351-360, 2015.

[212] E. Bazan, P. Dokladal, E. Dokladalova. Quantitative Analysis of Similarity Measures of Distributions. *British Machine Vision Conference*, 1-12, 2019.

[213] G Huang,., Z. Liu, K. Q. Weinberger,.L. van der Maaten, Densely connected convolutional networks. In CVPR, 2017.

[214] L. Zheng, Y. Yang and Q. Tian, "SIFT Meets CNN: A Decade Survey of Instance Retrieval," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 5, pp. 1224-1244, 1 May 2018.

[215] L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa, M. Pietikäinen, From BoW to

CNN: Two Decades of Texture Representation for Texture Classification. International Journal of Computer Vision, 127, 74-109, 2018.

[216] S. Jiang, J. Luo, GR. Pava, J. Hu, CL. Magee, A Convolutional Neural Network-Based Patent Image Retrieval Method for Design Ideation. Proceedings of the ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. Volume 9: 40th Computers and Information in Engineering Conference (CIE). Virtual, Online. August 17–19, 2020.

[217] I. Daubechies, et al., Ten lectures on wavelets, vol. 61, SIAM, 1992.

[218] Alot. ⟨https://aloi.science.uva.nl/public_alot/⟩.

[219] KTH TIPS. https://www.csc.kth.se/cvap/databases/kth-tips/download.html.

[220] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh. On the significance of real-world conditions for material classification. In European Conference on Computer Vision, volume 4, pages 253–266, 2004.

[221] E. Gomez, M. A. Gomez-Viilegas and J. M. Marin. A multivariate generalization of the power exponential family of distributions. *Communications in Statistics-Theory and Methods*, pp. 589 - 600, 1998.

[222] K. V. Mardia, J. T. Kent and J. M. Bibby. Multivariate Analysis. Probability and Mathematical Statistics Series. *Academic Press*, 1980.

[223] E. Gomez, M. A. Gomez-Viilegas and J. M. Marin. Multivariate exponential power distributions as mixtures of normal distributions with Bayesian applications. *Communications in Statistics-Theory and Methods*, pp. 972–985, 2008.

[224] J. K. Lindsey. Multivariate elliptically contoured distributions for repeated measurements. *Biometrics 55*, 1277–1280, 1999.

[225] J. Zhang and F. Liang. Robust clustering using exponential power mixtures. *Biometrics 66*, 1078–1086, 2010.

[226] N. Metropolis, A.Rosenbluth, M. Rosenbluth, A.Teller, and Teller, E. Equations of state calculations by fast computing machines. J. Chem. Phys., 21(6): 1087-1092, 1953.

[227] W. Hastings. Monte Carlo sampling methods using Markov chains and their application." Biometrika, 57: 97-109, 1970.

[228] J. Angulo, Morphological colour operators in totally ordered lattices based on distances: application to image filtering, enhancement and analysis, Comput. Vis. Image Underst.107, (1), pp. 56–73, 2007.

[229] M. Crosier, , L.D Griffin.: Texture classification with a dictionary of basic image features. In: Proc. Computer Vision and Pattern Recognition, 23-28, pp. 1–7, 2008.

[230] J. Zhang, M. Marszałek, S. Lazebnik, Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study. Int J Comput Vision 73, 213–238, 2007.

[231] S. Banerji, A.Verma, C. Liu. LBP and Color Descriptors for Image Classification. Springer-Verlag Berlin Heidelberg, 2012.

[232] M. Hassan and C. Bhagvati. Structural Similarity Measure for Color Images. International Journal of Computer Applications 43(14), pp. 7-12, 2012.

[233] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error measurement to structural similarity. *IEEE Trans. Image Processing,* 13(4):600-612, 2004.