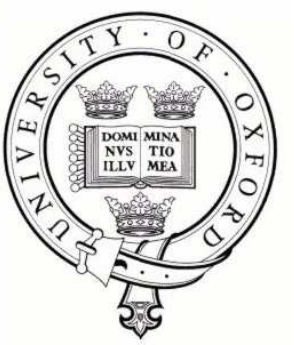
Structural Characterization of Liver Fibrosis in Magnetic Resonance Images



Anna Tünde Szilágyi

Supervisors
Prof. Sir. J. Michael Brady
Prof. J. Alison Noble
Dr. Catherine Kelly

Wolfson Medical Vision Laboratory
Department of Engineering Science
Trinity College
Trinity Term 2013

This thesis is dedicated to my brother *Ferenc*...

Abstract

The overall clinical motivation of this thesis is to differentiate between the different stages of liver disease stratifying into: no disease, mild disease, and severe fibrosis using Magnetic Resonance Imaging (MRI). As a related aim, we seek to differentiate as much as possible pericellular and non-pericellular fibrosis. This latter is clinically important, but currently no method exists that is able to perform this. Quickly, we realised that these aims push low level image analysis beyond their current bounds and so a great deal of the thesis is dedicated to extending such techniques before they can be applied.

To work on the most fundamental low level image analysis concepts and algorithms we choose one of the most recent developments, namely continuous intrinsic dimensionality (ciD), which allows the continuous classification of homogeneous patches from 1D structures to intrinsically 2D structures. We show that the current formalism has several fundamental limitations and we propose a number of developments to improve on these.

We re-evaluated feature energy statistics that were originally proposed in ciD, and additionally we examined the confidence one may have in state-of-the-art methods to estimate the orientation of features. We show that new statistical methods are required for feature energy, and that orientation predictability is more important than correctness of the estimation. This evaluation led us to the monogenic signal local orientation.

Analysis of feature or texture energy is also a main contribution of this thesis. Within this framework we propose the Riesz-weighted phase congruency model. This is able to detect internal texture structures but it is not capable of delineating boundaries. Nevertheless, it proves an appropriate basis for texture quantification.

Finally, we show that in contrast to using the standard established Kovesi approach, the developed texture measure leads to good results on the suboptimal T_1w MRI liver image staging images. We show that we are able to differentiate automatically between the separate disease scores and between pericellular and non-pericellular fibrosis.

Acknowledgment

I would like to thank to my supervisors Professors Sir J. Mike Brady, J. Alison Noble and Dr. Catherine Kelly for their support and friendship throughout this thesis.

The moral of my story is that I had decided not to go into the field of medical imaging until I met Mike on a usual Wednesday afternoon lecture at the Doctoral Training Centre and that changed my life in every aspect.

I could have never start in the Life Sciences Doctoral Training Centre without the support of Professor Simon Schultz and Dr. Marzena Wylezinska-Arridge from Imperial College London and my previous supervisor Professor Joseph Hajnal, Imperial College London. Further, my thesis would have been hard to continue without the support from Professor Tamás Roska, PPCU, Hungary.

Although it is not usual I am indebted and I would like to thank to my examiners their contribution: Prof. David Murray and Prof. Mads Nielsen who made very valuable comments and made my viva really enjoyable. I actually felt that it was too short and I would have carried on if possible.

I would like to thank to all of the people who have been working along me at the Department of Engineering, University of Oxford: Andrew McLennan, Dominique Van de Sompel, Candy Ho, Wang Po, Lydia Tanner, Ann Harvey, Niranjan Joshi, Mitchell Chen, Wen Jun Chi, Giselle Flaccavento and the list would continue if I would have endless space allocated, my special thanks go to Carolina Wessel.

Last but not least I would like to thank to my family: my siblings Ferenc, Ildikó and Gyöngyike and to my father to whom I am really grateful.

I am very grateful to the Engineering and Physical Sciences Research Council funding through the Doctoral Training Centre grant and their local coordinators.

Contents

Table 1: List of abbreviations used in the thesis.

Abbreviation	Explanation
MRI	Magnetic Resonance Imaging
RF	radiofrequency
FOV	Field of View
SE	Spin Echo
GRE	Gradient Echo
TE	Time-to-Echo
TR	Time-to-Repetition
PNS	Peripheral Nerv Stimulation
SNR	Signal-to-Noise ratio
PNSR	Peak Signal-to-Noise ratio
US	Ultrasound
CT	Computed Tomography
PET	Positron Emission Tomography
18-FDG	FluoroDeoxyGlucose
SPECT	Single Photon Emission Tomography
NHS	National Health System
HAI	Histology Activity Index
PDAC	Pancreatic Ductal Adenocarcinoma
cerulein	chemical compound to induce chronic pancreatitis
septa	fibrotic filaments in the liver
HVB, HVC	Hepatitis Virus B, C
HIV	Human Immunodeficiency Virus
ciD	continuous intrinsic dimensionality
PC	phase congruency
GET	Gradient Energy Tensor
BW_{RF}	bandwidth of the RF pulse
HOG	Histogram Oriented Gradient
LP, LO, LE	local phase, orientation, energy
fu, fv	Riesz components
OTPM	optimally topology preserving maps
PCA	principal component analysis
MLE	maximum likelihood estimate
HMMF	Hidden Markov Measure Field

Chapter 1

Introduction

This thesis aims to develop image analysis methods for the quantitative assessment and staging of liver disease based on Magnetic Resonance Imaging (MRI). As we describe, such an analysis necessitates: identification of the liver; extraction of the main lobes, vasculature, and the bile duct; and characterisation of the texture of the remaining liver tissue. From an image analysis standpoint, the liver boundary is mostly a step change in intensity, though the contrast with surrounding organs is often poor. The ducts and vessels are curvilinear structures whose width varies along their lengths, correspondingly changing from detection of the bounding steps to detection of a thin curve. This complex architecture is represented in Figure 1.1. The texture changes from a dense distribution of high frequency point-like features (early stage) to a mass of collagen links characteristic of late stage disease, Figure 2.9. This latter task is not straightforward even for a human observer, and to be robust and repeatable in the face of observed image changes from patient to patient, this set of challenges to automatic feature detection are currently beyond the state of the art, especially given the artifacts commonly found in liver MRI, which is taken in a single breath hold. For this reason, though our motivation is the staging of liver disease from MRI images, a great deal of the thesis is actually concerned with developing a novel,

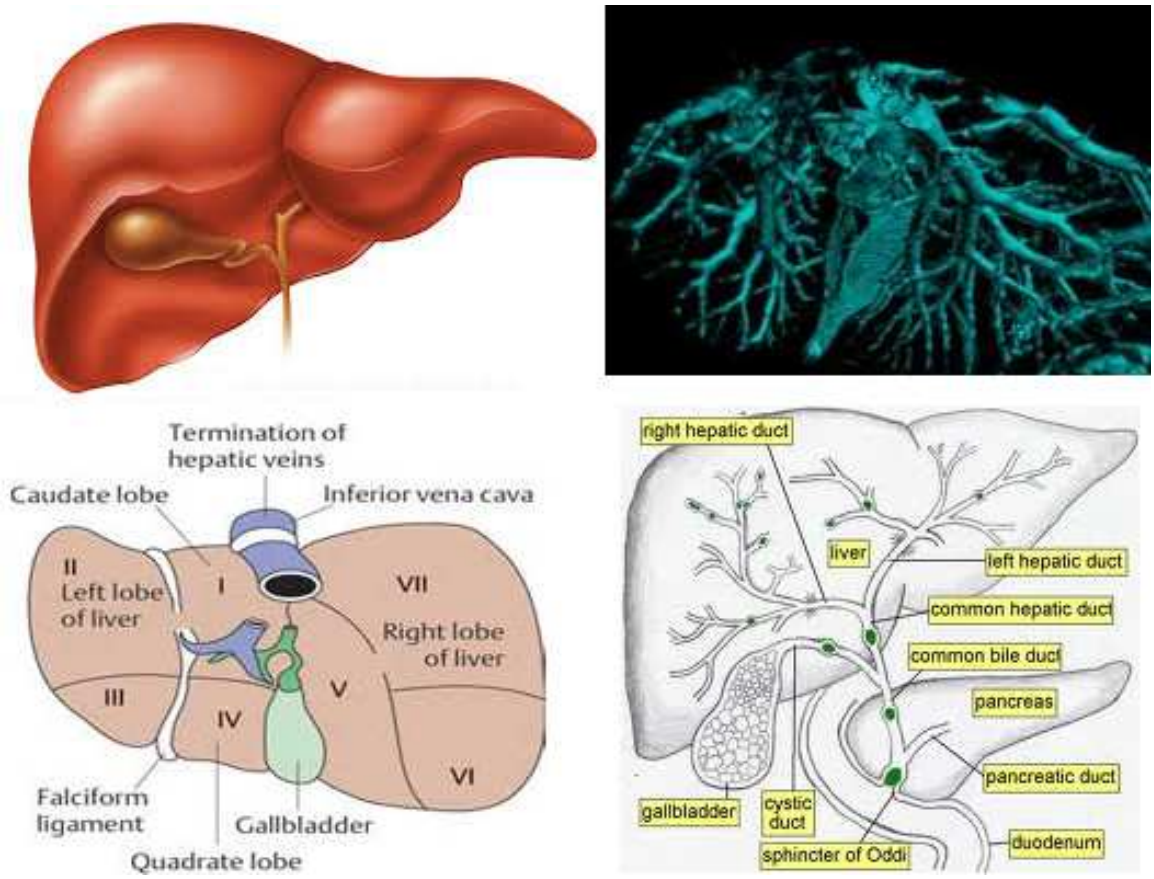


Figure 1.1: Complexity of the liver architecture including clockwise the liver (Centre, 2011), its vasculature (Institute image and video galery, 2009), duct system (Anatomy and System, 2013) and lobular system (MBBS Humanity First, 2011).

unified approach to feature detection, including a demonstration of the limitations of state-of-the-art methods. The method that we particularly selected is the phase congruency feature model. Therefore, the algorithms we develop have application beyond MRI images of the liver, though the final chapter returns to that principal application.

1.1 Problem formulation

Liver disease is one of the very few diseases whose mortality continues to rise with a constant rate year by year. Apart from biopsy, which is not just highly invasive but also samples only a small percentage of the liver, there is no current way to

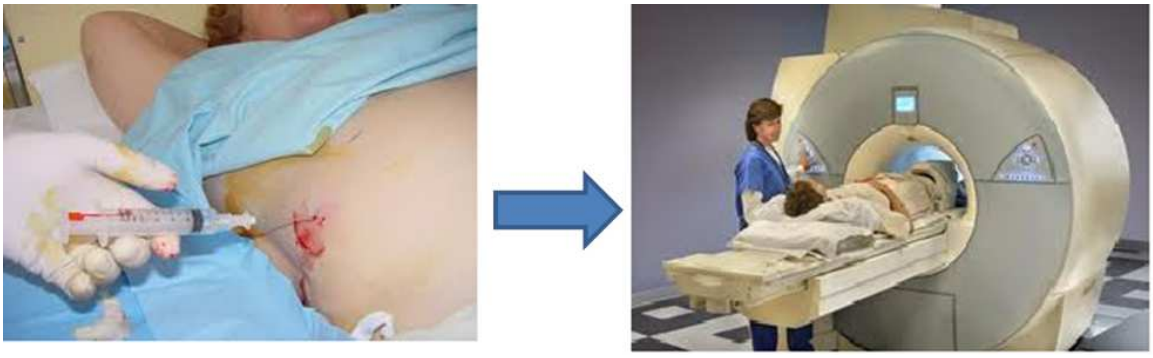


Figure 1.2: Transition from liver biopsy (Fierbinteanu-Braticevici, 2009) to MRI scan (Lefkovicz, 2013).

diagnose and follow-up liver patients. Our aim is to contribute towards providing a tool that allows this to be done quantitatively based on non-invasive measurements. One possible imaging modality that has the potential to do this, since it effectively has no depth restriction, gives excellent soft tissue contrast, and does not involve ionising radiation is Magnetic Resonance Imaging (MRI). This is illustrated in figure 1.2.

Viewed on an MRI image, liver disease changes the shape, the internal vessel structure and the tissue texture of the liver and poses a particularly hard image analysis problem both for shape and texture detection. It is therefore not surprising that there has been very little work on the image analysis to support hepatologists and their colleagues. The parts to be detected automatically vary enormously between patients and therefore a personalized approach is inevitable, further complicating image analysis.

We set out to focus on the texture changes typically seen during progression of the disease and as a result our first goal has been to examine the state-of-the-art in feature detection as a precursor to texture analysis. Our conclusion is that conventional computer vision approaches do not work well on these medical images.

Given the challenging task we are faced with, we experimented both with ideal features and a set of liver MR images, and whenever it is possible we show results on

both types of images.

1.2 Liver disease

The continuing rise in the incidence of liver disease is shown in Figure 1.3. This is in contrast to other causes, for example heart disease, cancer, stroke and respiratory diseases, which, in each case have seen a gradual slowing of mortality over the past two decades. Of course, these figures mask enormous variation between, say, breast cancer which has seen huge progress, and pancreatic cancer which has seen almost none. Liver disease mainly affects obese people, which is a rising proportion of the population throughout the developed world. For example, in 2010 the number of deaths from liver disease in the UK was 17000, with 25% of the population being obese and having the metabolic syndrome. The figures are even worse in the USA, where 5% of the population, corresponding to 15 million people, have chronic liver disease, and the metabolic syndrome and obesity affects 35%, which is currently around 100 million people.

The current reference to diagnose, stage, and follow-up patients is biopsy, and to a lesser extent hepatic venous pressure gradient (HVPG). Combined direct/indirect serum markers can avoid the need for liver biopsy in at most 50% of cases. Both biopsy and HVPG have a poor patient acceptance rate due to their highly invasive procedure and associated risk factors. The absence of robust noninvasive markers is the most significant barrier to patient management and design of clinical trials to test pharmacological treatment strategies. This thesis is a preliminary attempt to provide a noninvasive, semi-automatic method to perform the segregation of clinically valuable disease pathology.

In general, there are two main phases of liver disease progression which are dictated by their etiology. The two most common sources include viral hepatitis C (HVC, 170

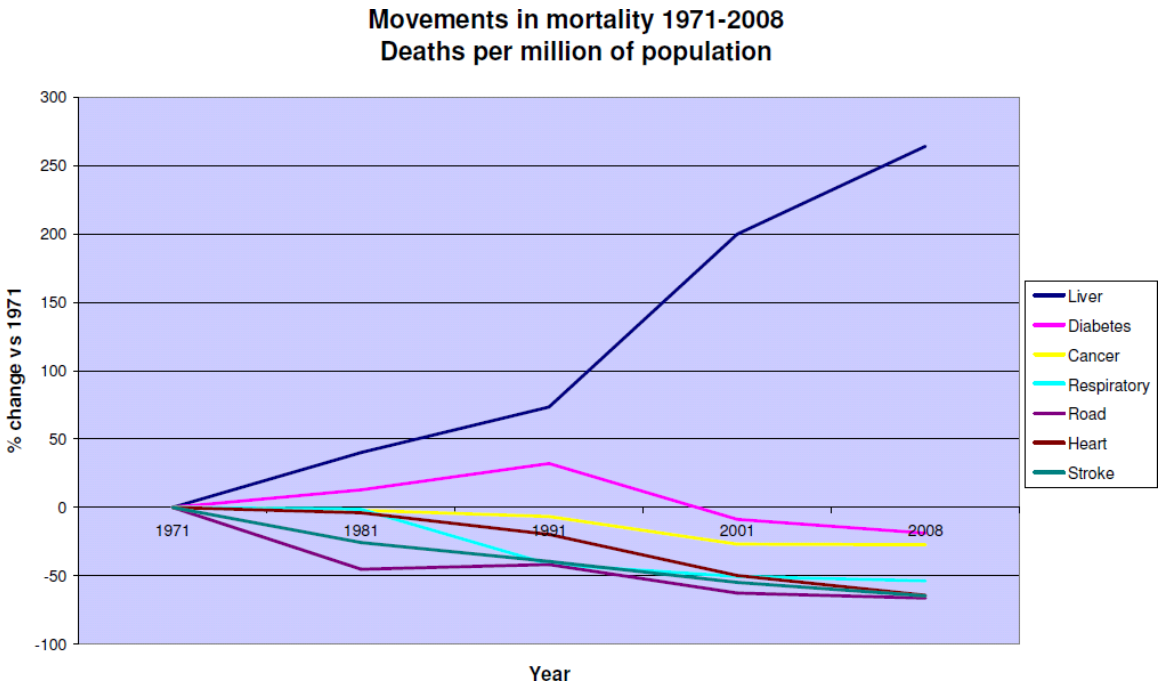


Figure 1.3: Major causes of death on the world (BLT, 2008).

million incidence worldwide/year, 350,000 deaths/year) and fatty liver disease (FLD), caused by obesity, poor diet, and alcohol misuse. Non-alcoholic FLD (NAFLD) affects 20-30% of all adults in the Western world (Day, 2011). It is diagnosed with >5% hepatic fat infiltrates. In addition, obesity related NAFLD is the leading cause of CLD in children. It is estimated that 17% of all children in the Western world are overweight. Of these, 70-80% have NAFLD (Giorgio et al., 2013). NAFLD progresses to the fibrosis/cirrhosis cascade through the intermediate stage of non-alcoholic steatohepatitis (NASH) following the "two-hit hypothesis".

Despite the fact that they follow different treatment regimes, the second stage of both diseases involves the development of fibrosis into cirrhosis and in 2-4% to hepatocellular carcinoma (HCC). Patients in the second stage experience a decrease in their quality of life and the necessity of liver transplant in progressive cirrhosis and HCC. In addition, the treatment constitutes a huge financial burden to healthcare systems.

Liver fibrosis is a result of chronic liver damage and is diagnosed by the accumu-

lation of extracellular matrix (ECM) proteins, with an increase in collagen, mainly associated with hepatic stellate cells (HSC) as a response to the presence of certain cytokines or mechanical/toxic stress. The excess of ECM proteins are often exploited in the indirect detection of fibrosis, for example using the chemical properties of proteoglycans. Current state-of-the-art results recognize fibrosis to be a reversible process which has significant consequences in patient management.

Biopsy is the current standard for the assessment of inflammation and fibrosis phases of progression, however it has several drawbacks. It is associated with 1-5% of hospitalization due to hemorrhage and infection, 0.57% of severe complications (Taouli et al., 2007), and mortality of 0.003% (Talwalkar et al., 2008). It has a reduced sample size (<25 mm) and portal tract number (<11) limiting accurate staging and grading due to sampling error of up-to 25-40% (Ratziu et al., 2005)[1]; it is subjective/qualitative which results in intra- and inter-observer variation of the histological interpretation; it is also difficult to repeat. Despite its drawbacks, it is the only current accepted method that is able to grade necroinflammation and stage fibrosis at the same time. Fibrosis histopathology differs both with the causative effect and the developmental stage of the liver (Giorgio et al., 2013).

Biopsy specimens are classified according to different scoring systems, all of which however have their own limitation[2]. The histology activity index (HAI), also known as the Knodell score, is based on necrosis, inflammation and fibrosis. The Ishak score is a modification of the HAI with the same inflammatory grades (0-18) but different fibrosis stages (0-6). The Metavir score is a modified version of the Knodell scoring system, but it was mainly verified for HVC related fibrosis stages (0-4). Despite the fact that histopathology of fibrosis for children aged 2-18 is different from that of adults, there are no current standard scoring systems accepted (Giorgio et al., 2013).

[1]Fibrosis is heterogeneous, especially in the clinically important early and intermediate stages of the disease.

[2]The current list does not include all the scoring systems.

Clinical trials to examine new treatment strategies against fibrosis have been challenged by two main factors. First, pre-clinical (rodent) models are of limited use because, possibly for genetic reasons, they do not develop fibrosis as a response to the same toxic effects as humans[3] (Starkel and Leclercq, 2011). Second, human clinical trials are cumbersome to design because of the practical feasibility of longitudinal disease monitoring. This necessitates the development of non-invasive methods, for example using medical imaging. This would improve the patient acceptance rate, allow longitudinal monitoring, and overcome sampling inaccuracies. The limitations of inter- and intra-observer variation could also be addressed using automated image analysis techniques. This thesis addresses one particular approach to this latter aspect.

Fibrosis is characterized by molecular, structural, textural and mechanical changes in the liver, and leads to progressive stiffness due to the collapse of liver parenchyma. These factors can be exploited by different *in vivo* imaging modalities. Molecular changes in humans are hard to monitor due to the complex vascular supply to the liver. Structural and textural changes in humans are observable both on ultrasound (US), Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) images. Severe liver damage, e.g. cirrhosis, is characterized by nodules smaller and larger than 0.3 cm (Macarini and Stoppino, 2013)[pg. 115], which is detected by all modalities though to variable extents. US and US based transient elastography, which quantifies shear wave velocity (FibroScan®, Echosens®), is only able to diagnose this very late stage, which has a quadratic change with fibrosis histology, and it has an imaging depth restriction, which further limits its applicability in the case of obese patients, who are, however, over-represented among patients. CT has no depth restrictions, but it is only sensitive to morphological changes which characterize the late CLD and

[3]Monkey and canine CLD and fibrosis models also exist, and follow common inflammation pathways with humans, especially in indigenous induction of the disease. They hold a great potential for imaging fibrosis, but current studies are limited.

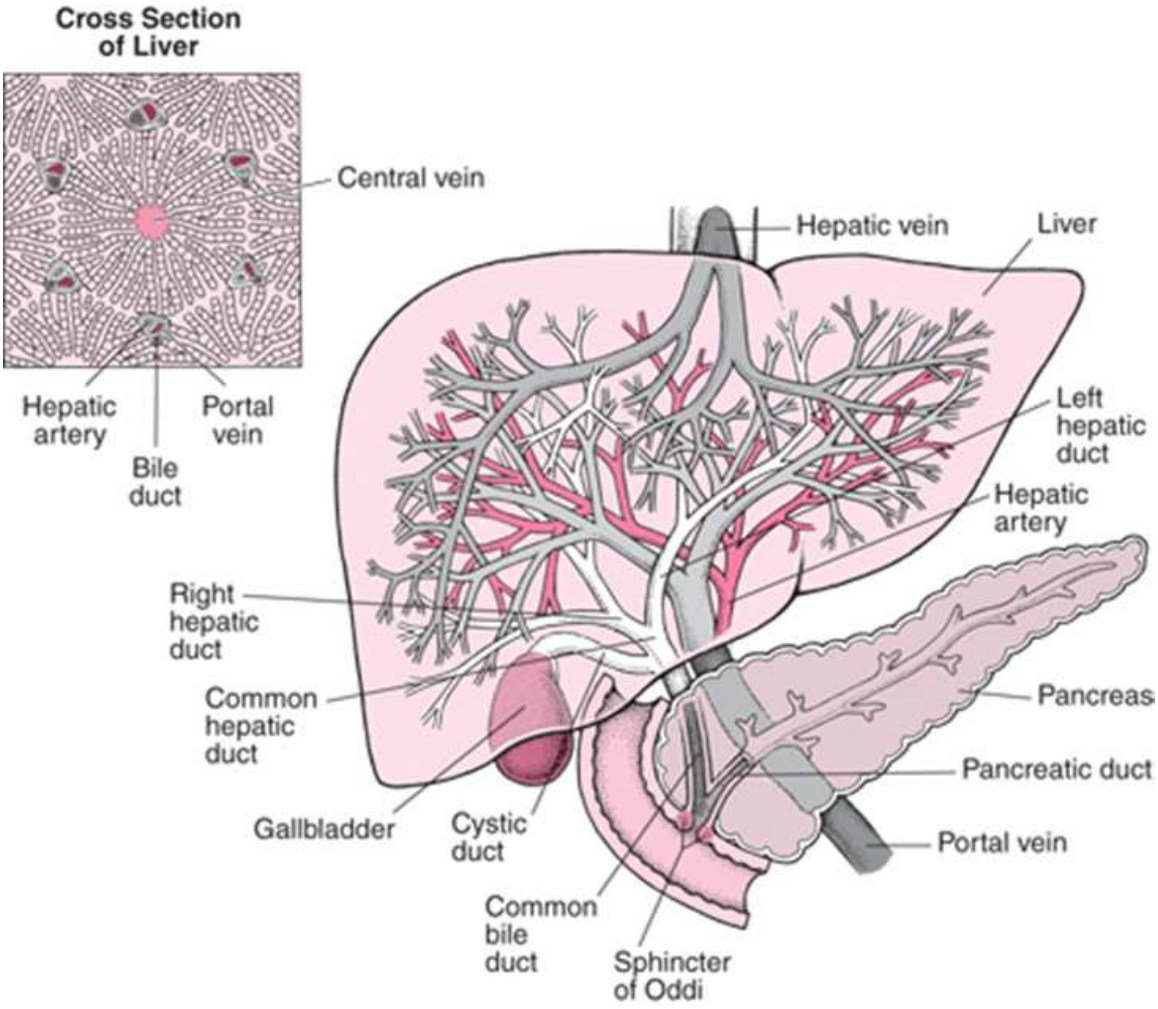


Figure 1.4: Liver vasculature and biological environment (Cohen, 2006).

its use for longitudinal assessment is limited by radiation safety. Magnetic Resonance Imaging (MRI) overcomes both the depth and radiation restriction and offers excellent soft-tissue contrast.

The liver is situated under the rib cage to the right hand side of the body and is composed of two main lobes: the right and the left lobe (there are classifications that are actually identify more than two lobes). To perform its function, the liver is composed of billions of liver cells that absorb and make available nutrients. The variety of functions that are performed by the liver is helped by the immense vasculature that perfuses it, Figure 1.4.

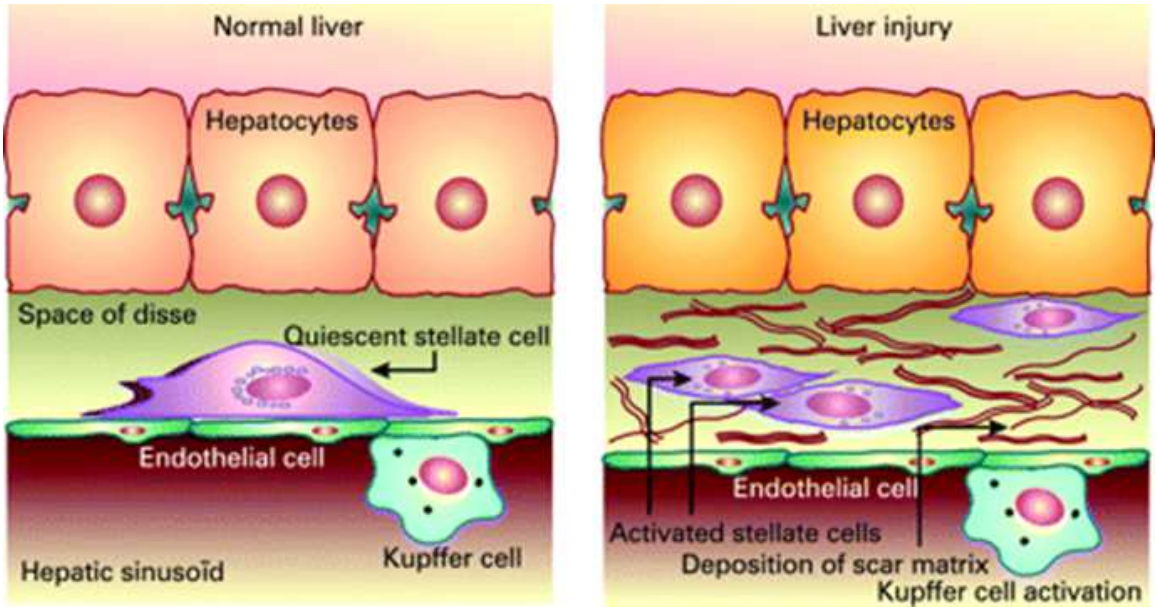


Figure 1.5: Hepatic stellate cells and their role in fibrosis formation (Asselah et al., 2009).

The two main blood vessels are the hepatic artery and the portal vein. The function of the liver is threefold: to *filter* waste, bacteria, and poison; to *process* nutrients (decomposed food); and to *store* fuel in our body. The blood delivers food particles in the form of proteins, fats and carbohydrates. This is how the rest of the body can process it in excess of immediate tissue requirement. The liver stores these, and in return makes vitamins, minerals and carbohydrates. The proteins that are broken down form new proteins that are then the basis for blood clotting and cholesterol.

The cellular level structure of the liver is illustrated in Figure 1.5. During normal functioning we have the following three layers: hepatic sinusoid with Kupffer cells, the Space of Disse with quiescent hepatic stellate cells divided by a layer of endothelial cells, and the hepatocytes.

During insult (which can be due to alcohol consumption, mechanical stress or unknown factors) the quiescent stellate cells acquire an active state during which they divide and promote the formation of fibrosis, Figure 1.5 right.

Until recently, it was believed that the formation of scar tissue is an irreversible process. Recently, it has been shown that if the source of the insult is stopped the process may revert back to a normal state without any residual scar tissue (Iredale, 2007), Figure 1.6. This recent finding changed current understanding and challenged sufficiently the community to put more effort into the research of formation and progression of fibrosis, without as yet any breakthrough.

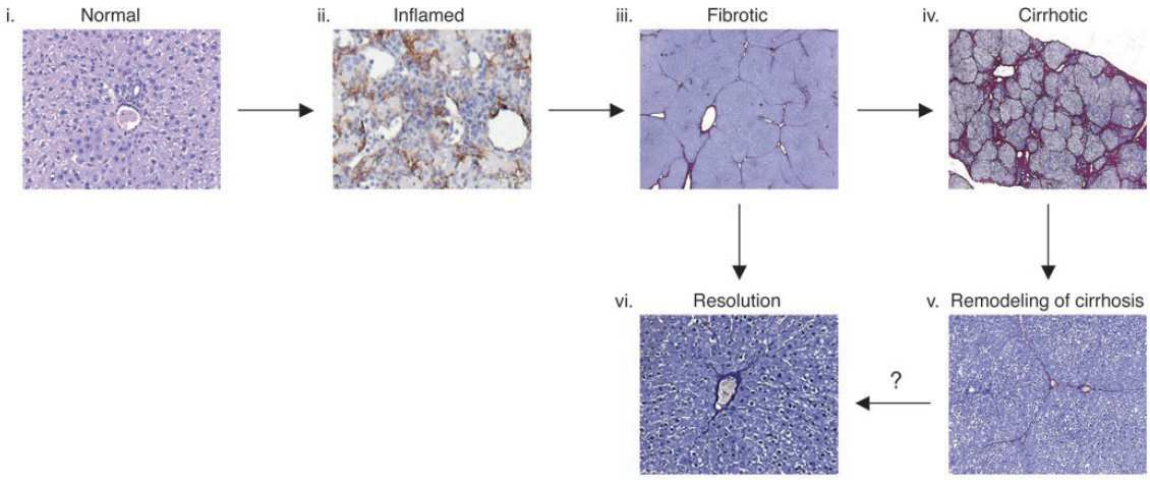


Figure 1.6: Resolution of liver fibrosis and cirrhosis. The figure was adapted without change and admission from (Iredale, 2007).

1.3 Human liver assessment in the clinic

Liver patients in the clinic typically present with abdominal discomfort that may be a sign of several diseases. The evaluation of the medical symptoms starts with an analysis of the vital functions from body fluids; for example, urine and blood, looking for specific markers. If this indicates anomalies, or the symptoms persist, then an abdominal scan is performed. Based on the necessarily restricted budget of the NHS, low cost scans are performed first. In practice, this means an ultrasound (US) scan followed by a computed tomography (CT), Magnetic Resonance Imaging (MRI) and finally Positron Emission Tomography (PET) if PET is indicated. All of these modalities reflect different properties of the tissue. The US and MRI are best for the assessment of soft-tissue, while CT will mainly be useful where density of the tissue is significant. A genuine MRI liver scan with a specific pulse sequence (for explanation see later) to detect iron concentration in the liver in ten minutes is shown in Figure 1.7.

Grading of liver disease in the clinic is a function of the extent of the apparent scarring and inflammation (Mazoff, 2010). There are three well accepted grading/staging systems, namely the Metavir, Knodell, and Ishak score systems.

In the Metavir system there is a difference between the grade and the stage of the liver disease. The grade means a score based on the amount of inflammation, where 0 means no activity and 3-4 means high inflammatory activity. The stage of the liver disease is assigned based on the extent of fibrosis between a score from 0 to 4:

- 0: means no scarring
- 1: means minimal scarring
- 2: scarring is present but it extends outside the liver and contains blood vessels
- 3: is bridging fibrosis that is spreading and connecting to other areas to contain

fibrosis

- 4: cirrhosis or advanced scarring of the liver

The Knodell system (Ishak et al., 1995) is slightly more complicated but some clinicians prefer its use due to thoroughness. The score itself is based around the Histologic Activity Index (HAI) (Ishak et al., 1995). In this system the following agreements exist:

- 0: no inflammation
- 1-4: minimal inflammation
- 5-8: mild inflammation
- 9-12: moderate inflammation
- 13-18: marked inflammation

where the first component indicates periportal and/or bridging necrosis, score 0-10, second is intralobular degeneration and portal inflammation score 0-4, and finally the combination of these indicate the amount of fibrosis with the fourth components meaning from no scarring to extensive scarring or cirrhosis (0 to 4).

The Ishak system (Ishak et al., 1995) is a modification of the Knodell scoring system above. The main driving force behind the Ishak score is to provide a measure of architectural changes, such as fibrosis and cirrhosis. This intent was clear from their early publication (Ishak et al., 1995):

- 0: No fibrosis
- 1: Fibrous expansion of some portal areas, with or without short fibrous septa
- 2: Fibrous expansion of most portal areas, with or without short fibrous septa

Figure 1.7: Liver Magnetic Resonance scan (FerriScan, 2009).

- 3: Fibrous expansion of most portal areas with occasional portal-to-portal bridging (P-P)
- 4: Fibrous expansion of portal areas with marked P-P and portal to central bridging (P-C)
- 5: Marked bridging (P-P or /and P-C) with occasional nodules (incomplete cirrhosis)
- 6: Probable or definite cirrhosis

The septa in the Ishak scoring means expansion of portal tract edges, without formation of bridges as most scoring system would identify.

Treatment is decided based on whether the Metavir score is greater than 2 or the Ishak/Knodell score is greater than or equal than 3.

The scoring system is chosen by the examining institution. In general, there is a choice between the Metavir and the Ishak score system. Many places consider Ishak scoring to be more reliable, because it gives a continuous measure of fibrosis on a 0-6

scale.

1.4 Technical challenges

One of the major challenges we face during this research is that the images are recorded with the end goal of a physician looking at major structures, not generating the best contrast of the liver parenchyma to enable for automatic texture analysis, despite the fact that texture is what clinicians look at in order to differentiate disease from non disease.

1.4.1 Feature extraction in human and mouse livers

Figure 1.8 compares a mouse and a human liver axial MRI image. They appear similar but at the same time different, not just because of the contrast. For example, the mouse sample lacks the veins and arteries that are clearly visible in the human scan. The base architecture of the liver shows up differently. Finally, despite the percentages of the liver on the images, we have a greater volume of liver tissue from

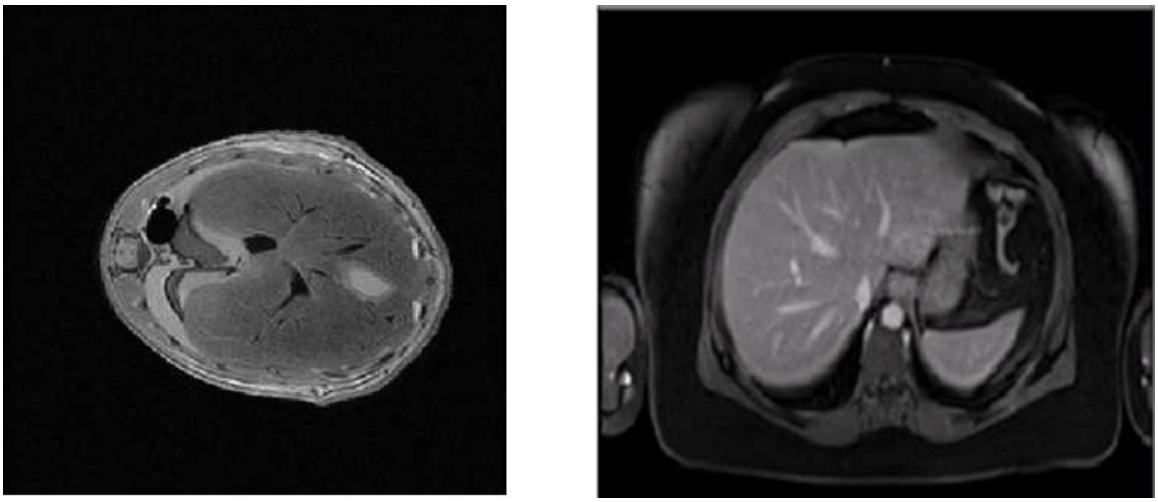


Figure 1.8: Magnetic Resonance Image of a mouse (left) and a human liver - axial slice shown. The mouse image was generated at the Small Imaging Laboratory at the Radiobiology Research Institute, Oxford, 2009, while the human scan was kindly provided by Acuitas Medical Ltd., Swansea.

humans than from mice. Because we are to detect tissue texture, the greater volume, i.e. human scans, are of greater use for us, despite the fact that they are of lower resolution.

1.4.2 Medical imaging - Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is in essence based on the assumption that it is possible to detect pathology based on the water content of tissues. Chapter 2 describes in more detail the basic operation of MRI for imaging a living system. Fundamentally, a body slice is excited with a radiofrequency (RF) pulse during the presence of slice select gradients: this is how we inject energy in-to the system. Then, depending on the particular imaging sequence, several pulses are played out that either operate in the xy plane, or along the z axis defined by the net magnetization direction (the main magnetic field strength defined by the position of the super-conducting coils). Finally, we describe the tissue parameters that govern how it returns to equilibrium. There are two main parameters: the longitudinal relaxation described by the T_1 tissue parameter in ms, and the transverse relaxation described by the T_2^* relaxation also described in ms.

In MRI, we label the patient direction according to the following system: from left to right; anterior-posterior; and superior-inferior. Then, depending on which plane we are in, we talk about axial, sagittal and coronal slice, see Figure 1.9. This is the convention that we will use throughout the rest of this thesis.

Patient and animal inflammation/fibrosis is best detected with targeted pulse sequences. Fibrosis is scar tissue and so is very dense. As a result, short T_2 tissue is the target for MRI imaging (Holmes and Bydder, 2005). Using the molecular mechanism of fibrosis with proteoglycans is also a possibility, with special Gadolinium contrast agents. This is used in the pulse sequence dGEMRIC (Tiderius et al., 2004).

Every imaging modality will to an extent detect fibrosis because of the physio-

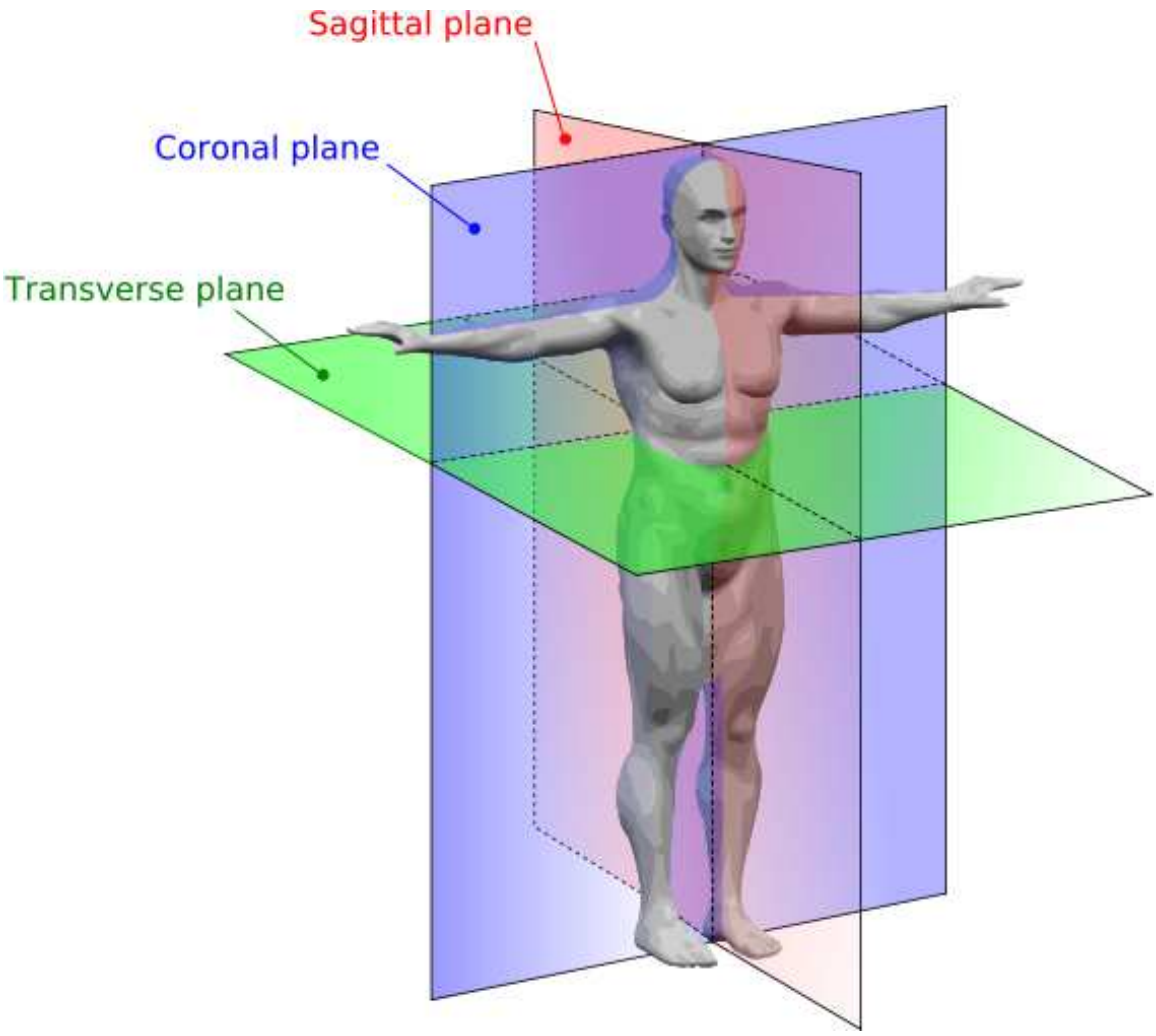


Figure 1.9: Magnetic Resonance Imaging: Imaging planes in the clinic (Mrabet, 2013).

logical changes that are associated with: tissue density, no blood flow, disruption to tissue architecture. However, not all of them are feasible in practice for the liver.

The method of imaging, however, will determine the image analysis methodology that can be used.

1.4.3 Phase and structure identification

Phase is the measure that allows one to compute qualitative measures of tissue structure and as a result its texture. It allows one to determine parity information on the

underlying signal irrespective of its orientation or local contrast. Phase based feature detection in the 1-D case was introduced using the analytic signal and the phase congruence (PC) model of features (Morrone et al., 1986), with the energy model (Morrone and Burr, 1988).

With the extension of the Hilbert transform into m-D, using the Riesz transform several signal models were proposed to derive the local phase. The 2-D monogenic signal uses spherical coordinates and geometric algebra to derive the local phase, energy and orientation; however the latter is only meaningful for i1D structures (Felsberg and Sommer, 2002). The method has been extended in two ways using two different signal models. First, the 2-D analytic signal (Wietzke et al., 2009) additionally allows an assessment of the apex angle, i.e. it characterizes an i2D event with an arbitrary angle between i1D events. Second, the conformal monogenic signal (Fleischmann et al., 2011) estimates curvature as the characteristic feature to the original monogenic features. The conformal monogenic signal is not included in our survey.

All of the monogenic based methods are based on differential all-pass filters. However, they perform much better in noisy situations than their structure tensor counterparts, though worse than integral invariant methods. We show that isotropic, model-free, scale space and curvature may be derived using integral invariants. However, due to their non-linear nature they have not yet gained much favour in the image processing community, and they will not be used in this thesis.

We conclude this chapter with the applicability of the above methods to i2D signals. It is really the integral invariants and the conformal monogenic signal model that allows the inclusion of i2D signals.

1.4.4 Shortcomings of the monogenic transform and advantages of the Riesz components for texture quantification

There are two main ideas that underpin the calculations in Chapter 4. One is the realization that the monogenic energy is of limited use even in the case of simple textured signals; the second is that the Riesz transform gives rise to a useful set of texture details.

The monogenic signal is a computational model that enables quantification of feature type and geometric orientation along feature energy. These are included in the local phase, orientation and energy of the monogenic signal of a certain image. As noted above, its calculation relies on the extension of the analytic signal into higher dimensions.

The Riesz transform is the natural extension of the Hilbert transform from 1-D into the multiple-dimensions using geometric or simple vector algebra (Nabighian, 1984) (Felsberg and Sommer, 2002). In the Fourier Transform (FT) domain, the Riesz transform is simply the isotropic extension of the *sign* function into higher dimensions. The monogenic signal is defined as the composition of the band passed signal and its two Riesz components. The two Riesz pairs are nothing other than a *sine* and *cosine* pair.

The image features are then calculated as follows: local energy, local phase and local orientation. These are described in detail in Chapter 3. In Chapter 4 we show that the Riesz components are more important in intrinsic texture description, or decomposition, as is local energy.

There were two driving forces behind the texture maps derived in Chapter 4. First, the Riesz components are powerful for texture characterisation, but they are not rotationally invariant; and second, the local phase of the monogenic signal is

rotationally invariant, and contains the second norm of the Riesz components; but it does not preserve all the qualities of the Riesz transform. This idea led us to the formulation of Riesz-weighted local phase and, later, Riesz-weighted phase congruency (PC).

This phase congruency is significantly different from the established energy weighted phase congruency. The Riesz-weighted phase congruency focuses on the details of the texture information, while neglecting the original ridge/edge, which are the main features that are detected by the energy weighted PC. Nevertheless, it is a powerful texture measure as is shown by our preliminary results on the different grade liver images and which are shown in the later chapters.

A measure that is rooted in statistics is the Fisher Z transform (Johnson, 2005) PC. The Fisher Z transform allows one to calculate cross-scale correlation and to carry out hypothesis testing on the correlation maps with pre-defined window sizes. We looked into three window sizes 5x5, 3x3 and 2x2. The drawback of the method is that it only allows pairwise cross-correlations to be calculated, so an extra step needs to be performed, the hypothesis testing on the gathered pairwise correlations to derive PC. These measures are based on the cross-scale correlations of the consecutive local phase maps from the monogenic signal.

1.4.5 Continuous intrinsic dimensionality

In Chapter 5 we provide a brief survey of the continuous intrinsic dimensionality that first gives a description of the properties of iDness, followed by a detailed discussion of the computation of intrinsic dimensionality proposed in (Felsberg et al., 2009).

We differentiate between 2D features based on their local environment. In this system we refer to homogeneous signals in which there are no changes, these are called the ci0D signals. Then we have the signals that have large variance in one direction but no change in the orthogonal direction, i.e. these correspond to lines and edges,

the ci1D signals. Finally, there are the signals that may show change in at least two directions, these are the ci2D signals.

We first introduce the preliminary models that laid down the foundations of the ciD from the Volterra filters. Then we present the proposed ciD system along the uncertainty sources that we envisage (Kalkan et al., 2007) at each mapping step. There is a thorough investigation regarding the mixture model of the noise and the squared gradient magnitude, as an image energy measure, which concludes that from a statistical point of view the Gamma distribution is best suited for this purpose.

Additionally, there is a statistical investigation with respect to the accuracy of the current state-of-the-art orientation estimation methodologies. We conclude that although the Gradient Energy Tensor (GET) gives the best result, the monogenic signal provides orientation that is the most consistent. This is followed by the preliminary ciD results in test images and compared to earlier methods.

Finally, we close the chapter by illustrating our proposed methodology.

1.4.6 Results - comparison between different stages of liver disease

In Chapter 6 we compare the fractal measure, entropy, clumpedness and Laws measure on phase based texture measure developed in this thesis and Kovesi's energy weighted phase congruency (Kovesi, 1999). Our quantification shows that all the presented methods produce statistically significant group median values. However, it is only the Riesz-weighted PC based Laws measure that also gives significant correlation with the seven scale Ishak scores.

Our developed phase based texture measure is shown to be an elegant solution for grading of different stages of liver disease and to differentiate between pericellular and non-pericellular fibrosis. Pericellular fibrosis is also called 'chicken wire fibrosis' because of its appearance; it first affects the central part of the liver before it invades

the rest of the lobules. This latter is clinically important information, no algorithm has previously been available to perform this. The method differentiates between no disease, mild and severe disease. Using fractal dimension we differentiated between pericellular and non-pericellular fibrosis. We close the thesis with these very promising results and a future work section with several possible new projects.

1.5 Conclusion

In this chapter, we have reviewed the biological motivation behind our technical developments, as well as the contents of this thesis. In particular, the definitions relevant to staging liver fibrosis are necessary to understand the results presented in Chapter 6.

The contributions of the thesis to the current understanding of our field are two-fold. First, we provide a technical local phase-based texture measure and, second, the application of the method in conjunction with well-established texture measures, such as Laws filtering provides statistically significant separation between three severity groups of liver disease. This is important because the measurements are based on suboptimal non-contrast enhanced liver MRI images. In addition, our preliminary experiments show that the method can differentiate pericellular fibrosis from the same Ishak score, which has not previously been reported.

Chapter 2

Medical Image Acquisition

There are several ways by which one may examine the source of a set of medical symptoms. In this chapter, we examine methodologies that allow one to do this non-invasively.

2.1 How can we look inside the body?

The current gold standard to assess fibrosis of the liver is by biopsy, of which there are 650,000 taken around the world each year. A biopsy involves piercing the patient's body several times with a 20 cm needle, collecting tissue which is then examined under a microscope. This is limited in coverage; indeed it examines approximately 0.002% of the liver! Apart from being hit-or-miss, providing limited information about the true state of the liver, the procedure has side effects resulting in possible major bleeding and/or long term injury to the organ. On the top of this, it is prone to inter-observer variability. Biopsies cost up to £600, with 2% having complications that cost up to £2000.

The shortcomings of biopsies cannot be alleviated by serum tests, because these do not carry spatial information. Furthermore, they can be completely normal even for severe liver disease cases.

All the above drawbacks of the current methodology indicate the use of imaging modalities that can estimate three major clinical scores: the extent of fibrosis, the fat content, and the iron content.

Medical imaging is necessary because there is no other non-invasive way that we can observe the inside of the body. The mode of imaging in the UK is largely dictated by the funding available to the National Health System (NHS), the task to be examined, and the possible side effects. The medical imaging systems therefore may be classified according to at least two criteria. One is the possibility of observing certain biological phenomena in the body, the second is the cost, which dictates how many people could access the systems.

From Ultrasound (US) to Positron Emission Tomography (PET) we have different scales on which the biological system may be observed. US detects the organ boundaries as impedances, and relies on the workings of a transducer that is made of a piezzo-electric element and is able to transform electrical signals into mechanical sound waves conversely. It does not use ionising radiation and is widely used in foetal and cardiac imaging.

Although US can assess liver size, blood flow and even in certain limited cases identify fatty liver disease, it is highly dependent on the depth of tissue to be examined. Many liver patients are obese *de facto*, and as a result US cannot reliably provide information on the extent of fibrosis, inflammation or iron content. Newer techniques such as US elastography may differentiate between normal and severe liver disease, however the depth limitations remain, and this technique cannot assess the whole liver and remains as a research tool.

Computed tomography uses ionising x-radiation to image anatomy. It is a great source of information in detecting bone structure, and given its lower price, it is the preferred modality to monitor treatment response for cancers. This is possible because CT measures radiological density and tumours are radiologically much denser than

any of the surrounding tissues. Tumours show up bright because they have higher absorption. Contrast agents also exist for CT, but are mainly used for angiograms, i.e. when one is interested in monitoring the defects of the vascular system. Another use of CT is planning radiotherapy, because it gives excellent anatomical definition and provides the necessary organ information for shaped beam, conformal radiotherapy.

However, even contrast-enhanced CT cannot identify diffuse fibrosis and, because of radiation burden, it cannot be used frequently to monitor therapy and disease progression. There is also a risk of cancer development due to its invasive ionising radiation, especially for young patients.

Magnetic Resonance Imaging (MRI) uses water content (actually, the presence of hydrogen nuclei) to detect pathology. It provides soft tissue contrast as compared to dense tissue contrast in CT, and it does not rely on depth as US does. Water relaxation as a result of radiofrequency (RF) excitation is at the heart of the method. It assumes that changes in water content or water flow is indicative of pathological changes in tissue. It is non invasive and although MRI contrast agents exist, they will not be used in this thesis.

In theory, an infinite number of contrast possibilities exist with the tuning of the parameters in the imaging sequences. It is the ideal methodology for a set of experiments because it is non ionising although the resolution of MRI is much lower compared to the anatomy definition available with CT. MRI is a superb functional imaging modality, even though, it is currently not possible to detect molecular changes with it.

PET and Single Photon Emission Tomography (SPECT) are the preferred ways of detecting molecular (metabolic) changes within a tissue. SPECT is used almost exclusively for cardiology. The most common radiopharmaceutical used with PET is 18-FDG (FluoroDeoxyGlucose), a glucose analogue. Although it is possible to detect inflammation with it (Signore et al., 2010), the kinetics of the 18-FDG tracer are

extremely complex because of the dual blood flow in the liver. For this reason, PET is not currently much used for liver assessment.

US allows real-time observation of tissue (although MRI and CT are able to achieve near real-time fast dynamic imaging using contrast agent), while an MRI takes anywhere between 15 min to hours depending on the pulse sequences deployment and the magnetic field strength. PET and SPECT acquisitions in fact typically start at 40 min after radiochemical injection and we get low spatial resolution.

Though cost is a major consideration in providing healthcare, it is less of a consideration in research, where the precision of the results is more important. That is why our initial US experiments were replaced by MRI scans, which provide greater functional details. An illustration of the MRI system is shown in Figure 2.1. In general, MRI is used with a T_2 spin echo sequence to examine the liver duct, but it does not genuinely provide information of fat content, fibrosis or iron content.

MR elastography uses non-invasive mechanical waves to evaluate the underlying tissue stiffness. Its main application was suggested for the assessment of liver fibrosis, where disease progression is marked by the accumulation of collagen, i.e. increase proportion and relative magnitude of stiffness. At present there is limited information about which of these two measurements correlate with the biological disease state.

2.2 Basics of MRI

In this section we review the basics of MRI, both from the point of view of living tissue examination and the hardware requirements that make these measurements possible.

The base assumption of any MRI examination is that pathology changes the water (more precisely the availability of spins of hydrogen protons) content of the system that is scanned, and, as a result, it is observable on the images produced. The spin is

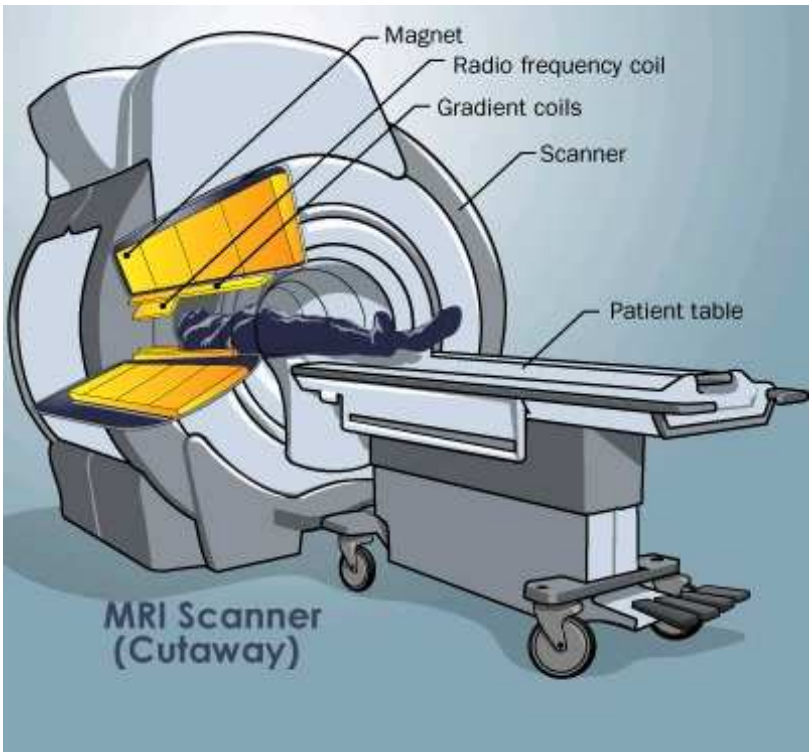


Figure 2.1: What the Magnetic Resonance Imaging (MRI) system is composed of? (Gould, 2013)

a tiny magnetic dipole moment generated by the rotating charged protons that come from the water molecules H_2 stripped off their electrons by the much stronger oxygen molecule. The available hydrogen spins of the tissues may be changing in different ways: either the tissue changes density and therefore its water content, or what is the same in fact changes direction of flow due to obstruction. Both result in an MRI signal that is different from the normal.

2.2.1 Essential concepts of MRI image formation

Like any imaging modality, MRI is composed of an excitation stage and a relaxation stage. During the excitation stage, energy is introduced into the system in the form of radiofrequency (RF) pulses, and by relaxation we mean the return of the excited body slab into its normal state. The way in which it returns enables clinicians to gain insight into the underlying tissue structures, including in many cases, disease.

In standard MRI, a single slice is excited using an RF pulse that is transmitted during the application of a linearly varying magnetic field ("slice select gradients"), see Figure 2.2, which is orthogonal to the plane to be imaged. The gradient changes the Larmor frequency of the spins. The base equation of precession is:

$$\omega_0 = \gamma B_0 \tag{2.1}$$

where ω_0 is the angular Larmor frequency of the spins, γ is the gyromagnetic ratio which for protons equals $2.7 \cdot 10^8$ rad s^{-1} T^{-1} and B_0 is the main magnetic static field. The linear gradient is used to be able to selectively excite only a slice of the body. The displacement the gradient makes in slice position is defined by (Delakis, 2006):

$$\delta z = \frac{BW_{RF} 2\pi}{\gamma G_z} \tag{2.2}$$

where G_z is gradient strength and BW_{RF} is the bandwidth of the RF pulse. This shows that the larger the raise in the gradient the thinner slice can be excited. We never achieve perfect slice selection because there is a range of frequencies that are present in the RF pulse. The ideal pulse in the time domain would be an infinite *sinc* because its Fourier Transform (FT) is a rectangle, hence perfect uniform bandpass. As a result, the perfection of the homogeneity of the slice selection depends how long we could play out the slice select pulse and ramp up the gradient. Then, the excitation is achieved for the spins that fall into the frequency range of the transmitted RF pulse (Figure 2.2).

The directionality of the spins varies within the body. They are randomly oriented in steady state (also zero field), so the overall magnetic dipole moment field they comprise is zero. However, when the same spins are placed in a sufficiently powerful magnetic field, they align with or against it. The difference in the numbers aligned

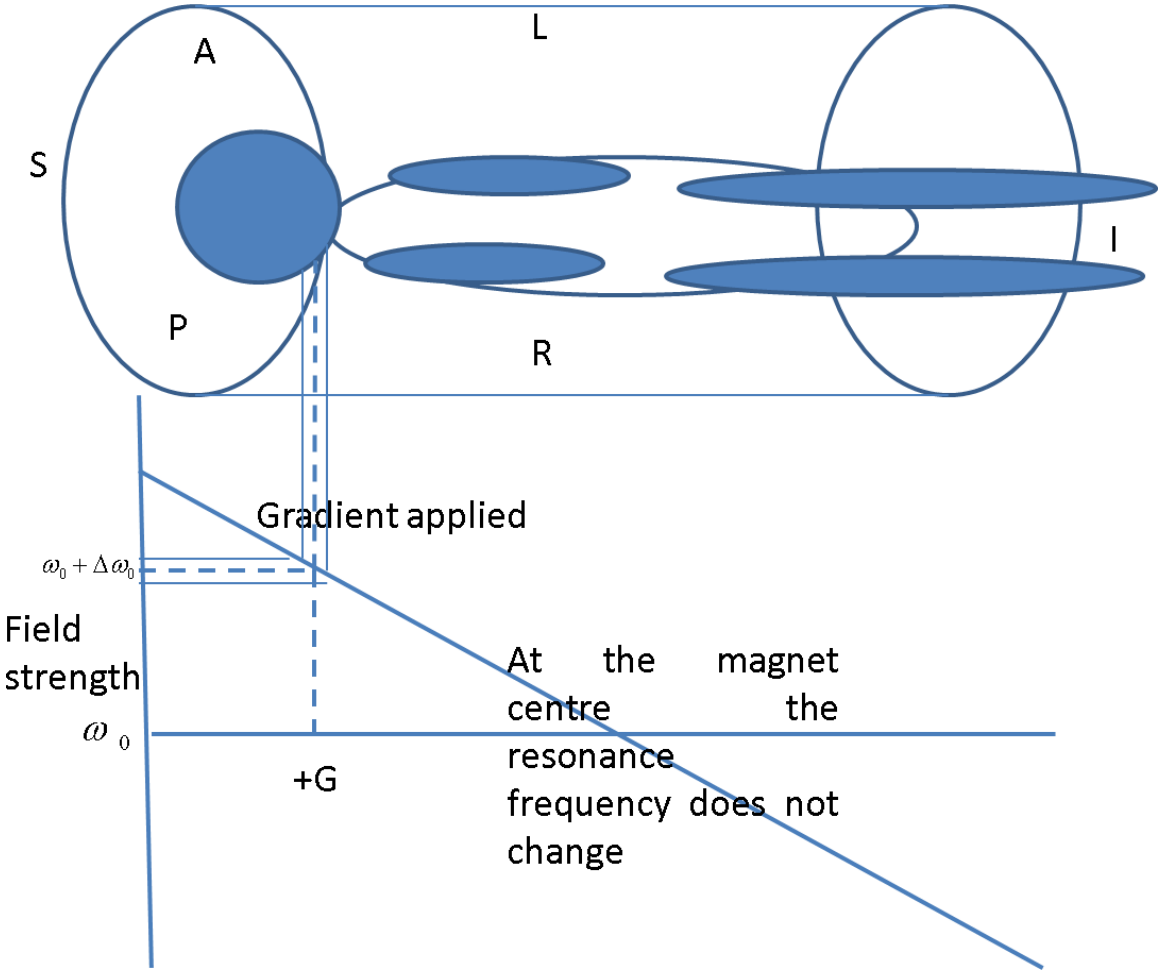


Figure 2.2: Slice selection using linear gradients. Notation: A is anterior, P is posterior, R is right, L is left, S is superior and I denotes inferior. The gradient G linearly changes the magnetic field along the main magnetization. This allows to select a slice with a centre frequency ω_0 and bandwidth $[\omega_0 - \Delta\omega, \omega_0 + \Delta\omega]$ which then together dictates the slice position and its thickness.

in these two opposing directions is proportional with the proton density:

$$N\uparrow - N\downarrow \approx \mathrm{PD}\frac{\gamma\hbar B_0}{kT} \tag{2.3}$$

where PD is the proton density, γ is the gyromagnetic ratio, B_0 is the main magnetic field in Tesla, k is the Bolzmann constant, $\hbar$ is the Planck constant, and T is the temperature. A (slight) majority of the spins are aligned with the main magnetic field. This is shown in the left hand side of Figure 2.3. First, it shows that spins

are both aligned and reversed to the z axis. The next figure reflects on the fact that we may do a vector addition, and get the transverse (xy direction) and longitudinal component of it (z component). The right hand side of the figure shows the relaxation phenomenon. First the spins are put down say with a 90 degree pulse to the xy axis. Then these start to precess and relax back to their original direction, this is described by the negative of a negative exponential function as a function of tissue parameters.

One way to express this phenomenon was described in (Bloch, 1946) as:

$$M_z = M_{z0}\left(1 - \exp\left(-\frac{t}{T_1}\right)\right) \tag{2.4}$$

where M_z is the overall net magnetization, t denotes time and T_1 is the spin-lattice relaxation time which depends on tissue properties. Different tissues have different longitudinal relaxation constants. They would show up with different slopes on the right hand side graph.

After slice selection and excitation, dependent on the particular pulse sequence, several pulses are played out that achieve the signal that is desired. Readout gra-

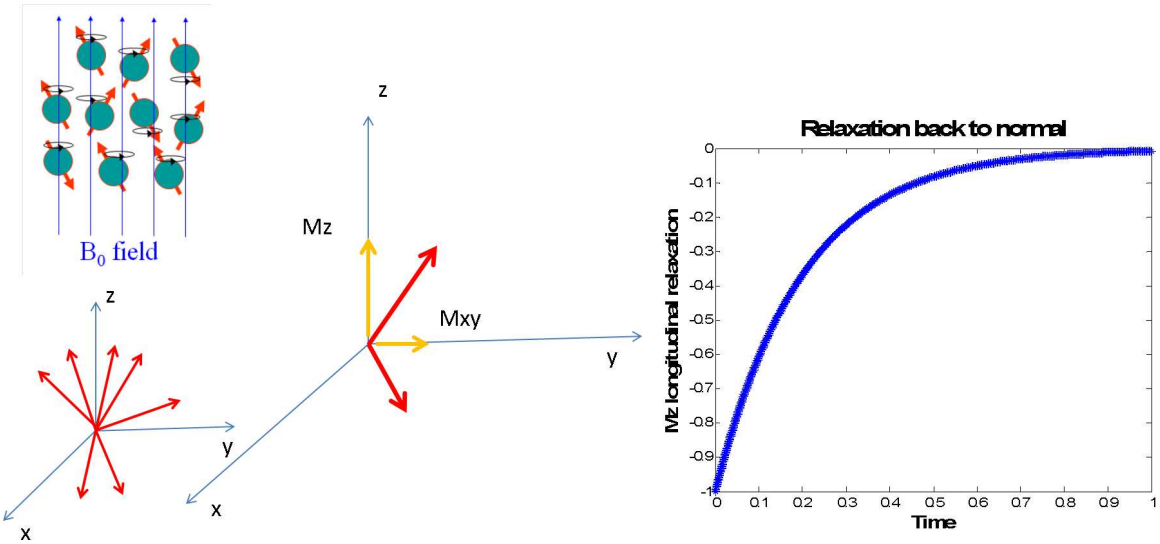


Figure 2.3: We excite in the state where the spins are mostly aligned with the B_0 field (illustrated in two ways), then we observe the longitudinal relaxation back of the main magnetization.

dients, which change the magnetic field strength linearly, are then applied when the echo is recorded into the k-space. The k-space represents the spatial frequency information of 2-D and 3-D objects and the location of each data-point is dictated by the gradients applied. Each and every k-space data-point, which represents the raw magnetic resonance response, in which the image information is recorded contains information about the whole slice. It is by applying Fourier Transform (FT) that the k-space raw data is converted into images that are read by the clinicians.

Figure 2.4 shows the original input image with its corresponding FT Fig.(a) and (b), then it shows that the missing centre part of k-space results in an image that highlights high frequency (edge) information Fig. (c), finally, it is demonstrated that padding with zeros the outside of the centre part gets rid of the above well defined edges and results in a blurred image. The k-space and the low frequency (blurred) image is shown in Figure (e)-(f).

There is a relationship between the field of view (FOV) and the extent of k-space that is recorded, also there is an inverse relationship between the spatial frequency and the space-domain inter-sample distances. One cannot arbitrarily undersample or reduce the FOV in the phase encode direction (this is the time costly step) because this results in wrap-around effects, or folding of the image. Different pulse sequences are designed for different pathologies.

2.2.2 MR parameters for tissue characterization

Watery materials have very long T_1 values because the environment vibrates faster than the Larmor frequency and it is harder to unload the extra energy, in the more solid tissues however the molecules' neighbourhood vibrates closer to the Larmor frequency and therefore we have short T_1 values. Tumours have a high T_1 due to the amount of inflammation (water content) that characterizes them.

There is another tissue property which is called the T_2^* relaxation or spin-to-spin

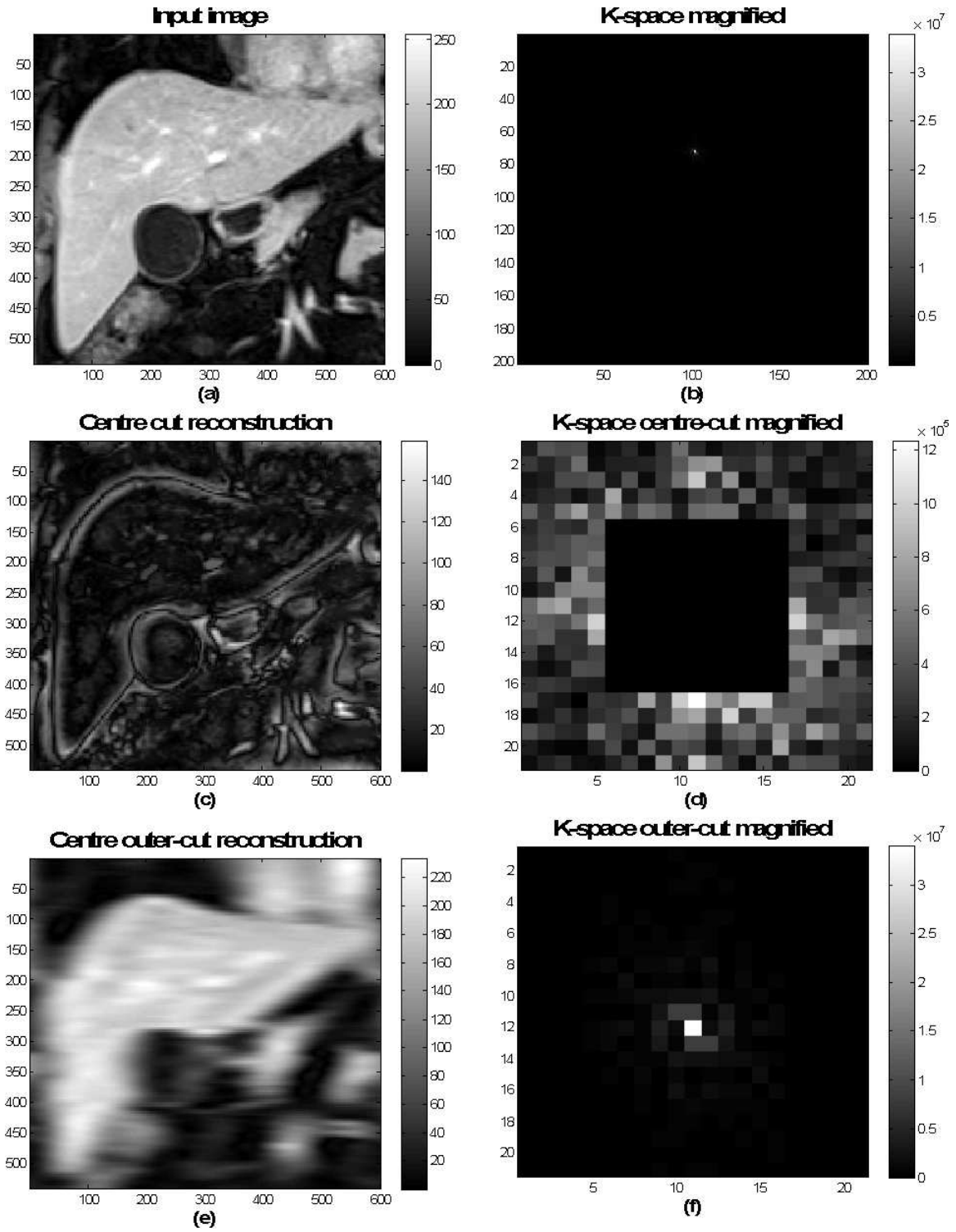


Figure 2.4: Effect of k-space points at different positions within the map.

relaxation time. We measure it by the decrease in the signal in the xy direction

(Delakis, 2006):

$$M_{xy} = M_{xy0}\exp\left(-\frac{t}{T_2^*}\right) \tag{2.5}$$

here M_{xy0} is the transverse magnetization immediately after the application of the RF.

Liver MR images are recorded during a breath hold to avoid motion artifacts that would result in ghosting from the fact that the k-space points are recorded in the wrong spatial frequency place. This is the way that the images analyzed in the end of this thesis are recorded. Unfortunately, the MR images that we have are of very low quality despite the fact that they have been taken during a breath hold.

2.2.3 MRI pulse sequences

The simplest pulse sequences are the gradient echo (GRE) and the spin echo (SE). We briefly describe how to read a pulse sequence in order to understand what the previous sentence means, then we describe in more detail the liver acquisition sequences (Figure 2.5-2.6).

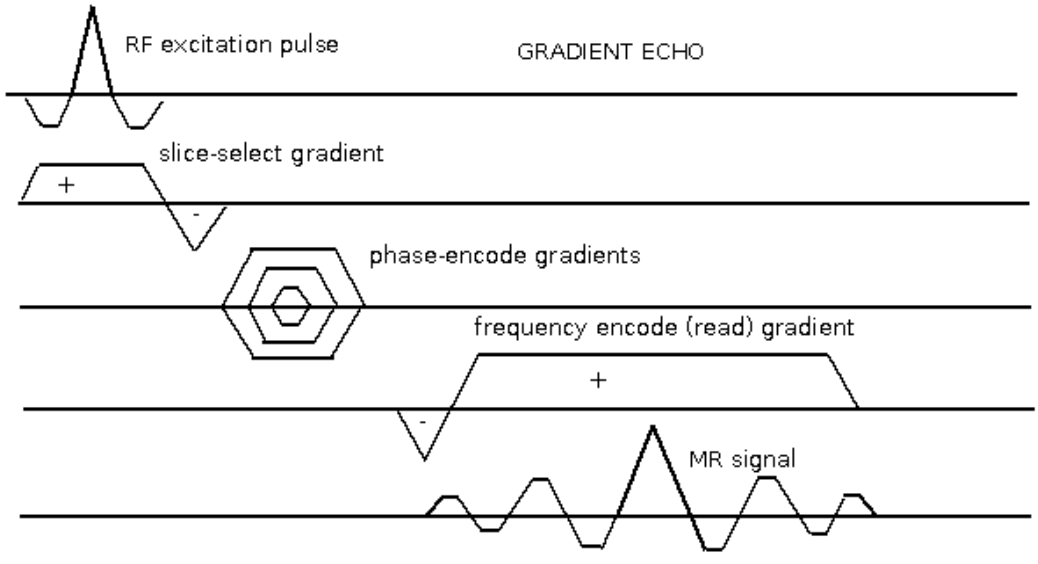


Figure 2.5: Gradient Echo pulse sequence for MR imaging.

First, examine Figure 2.5, which shows a gradient echo sequence. The pulse sequence needs to be read from left to right, corresponding to the time flow. As described earlier, an RF pulse of 90 degrees is played out along the application of the slice select gradients. This will tip all the spins onto the xy axis. The negative part of the slice select gradient is there because some of the spins start to dephase and this rephases them. Two things then happen: the spins start to dephase in the xy plane and build up a component along the z axis. This is followed by the phase encode gradients signalled by the different height parabolae. The different pieces signify different lines in the k-space; this is the time consuming encoding step. After a negative re-focusing gradient we follow the read-out gradients during which time the signal echo is sampled (MR signal on the figure). The figure also shows that we measure time to echo (TE) from the centre of the 90 degree pulse and the centre of the recoded MR pulse. In turn, time to repetition (TR) signifies the time spent between two 90 degree pulses. The process is followed by the exact same procedure.

There are several differences and similarities between the gradient echo and the spin echo sequences. Figure 2.6 shows the spin echo sequence. Here we have a second pulse that can achieve a stronger signal, as follows. A 90 degree pulse is played out during the slice select gradient followed by refocusing and phase encoding. This means that all the spins are tipped now onto the xy axis and we have faster and slower precessing spins depending on the RF pulse and gradient during excitation. To achieve a stronger signal all the spins are tipped by a 180 degree pulse onto the opposite side of the xy axis. This means that all spins will again be in phase after a certain amount of dephasing time.

Why will all the spins be in phase after the 180 degree pulse? Because the slower spins will be tipped closer to the opposite side and the faster spins will be tipped further away. Since they follow the same path, they will arrive in the same location in the same time. This is when imaging takes place. We have a read-out gradient

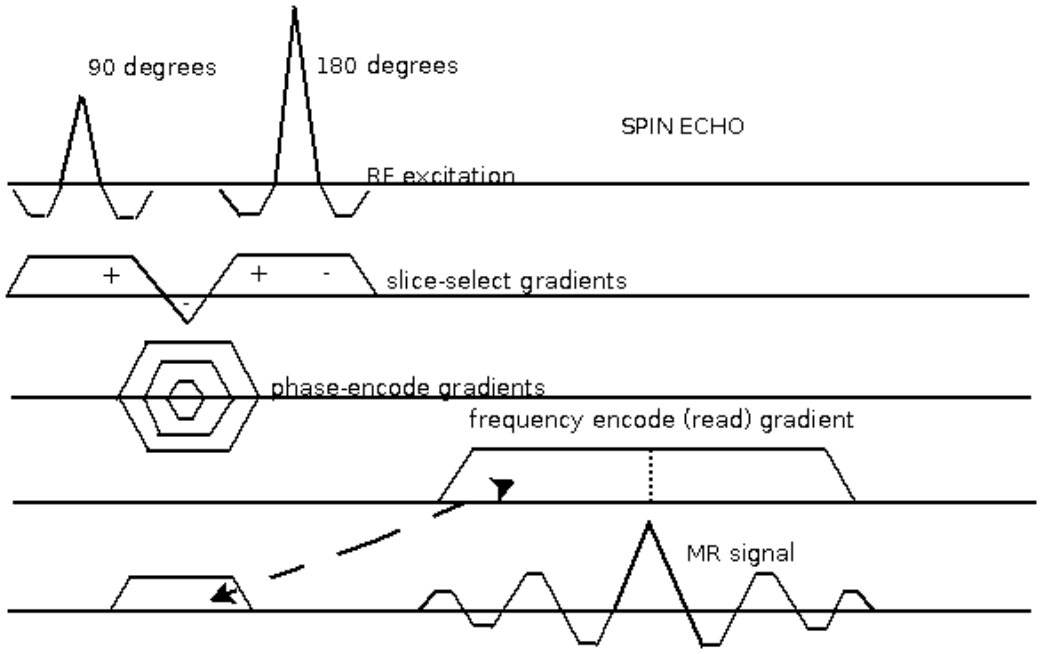


Figure 2.6: Spin Echo pulse sequence for MR imaging.

(there is no need for re-focusing this time) during which the MR signal is recorded. We have the exact same meaning for the time-to-echo (TE) and time-to-repetition (TR) as before. This process repeats until all the k-space points are filled.

2.2.4 Hardware requirements

To understand how the measurements are really made we need to first illustrate the main blocks of the MRI hardware (NessAiver, 1997).

A typical system is shown in Figure 2.1 with a patient already lying in the scanner. The main parts of the system are highlighted and named as the coil system and the scanner itself with the sliding table. A simplified skeleton version of this image may be found in Figure 2.7. This shows how the excitation and slice selection coils are packed into one system. The shimming coils and the computer control over the gradient adjustment and the RF coils along the display settings are also shown. In reality this is the structure of the magnet coils like an onion system from inside out:

1. RF coils

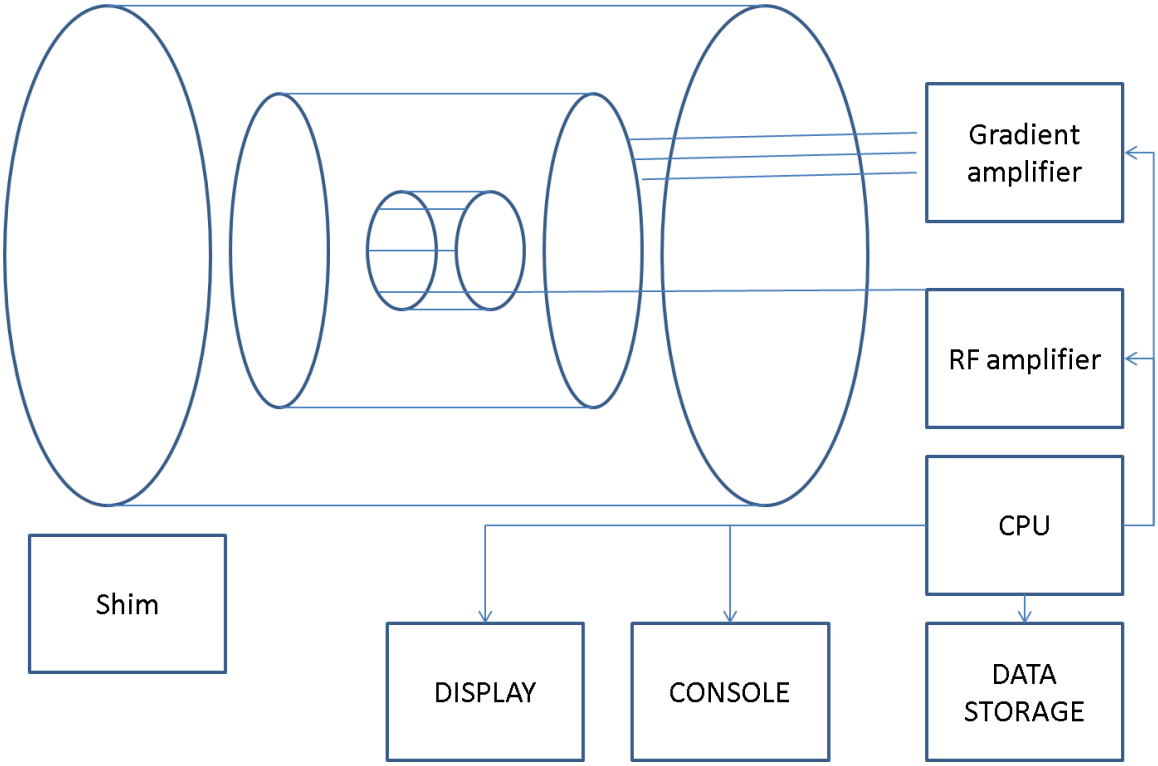


Figure 2.7: Magnetic Resonance Imaging (MRI) hardware appliances.

2. gradient coils

3. shimming coils, and

4. superconducting coils bathed in liquid Helium

The vast majority of current high field ($\geq$1T) MRI systems are superconducting, which means that current flows in wires that have no resistance. To achieve this the wire has to be cooled below its critical temperature, which is typically a few degrees Kelvin. This is why the cooling system is required and brings a higher maintenance cost to it than general imaging systems. Liquid Helium is currently used for the cooling system to maintain this state.

This figure contains the shimming coils. Shimming is required because high field magnets inevitably have inhomogeneities. When we have a separate coil system we talk about active shimming, or multiple electromagnets that are used to shape little

parts of the current flowing through the wires. It is slightly harder to control this from behind the computer bore and more expensive to actually make them. One must be careful, as opposed to field homogeneity passively shielding the magnets is also used to prevent the nearby objects from the exposure of fringe field. This does mainly add to the weight of the system because a layer of iron is added to the magnet. To avoid the fringe field one might also do it actively by introducing secondary coils opposite to that of the primary coil system.

The gradient coils are used to ramp up and down the magnetic field with respect to the Larmor frequency to be able to select a slice for excitation and readout. The gradient coils cause both a temporal and linear change in the main magnetic field strength. There are three gradient systems, one for each of the three dimensions, which are the three orthogonal directions x-y-z. Gradient coils typically come in pairs to be able to generate linear field strength. In their original form the gradient coils have a Gaussian profile and two coils with opposing currents are able to generate a linear gradient around the main magnetic field. Note, that the field will only be linear around this point. By rule of thumb the frequency at the centre of the magnet never changes (NessAiver, 1997). Once we have this three directional linear system we can navigate within the body under examination.

The characteristics of the gradient system determine the type of imaging that may be performed with the given system. The speed of imaging is dictated by the so called slew rate and the amplitude of the gradient. The slew rate refers to the rate of change of the gradients. The amplitude in turn means how high a gradient can be produced and is described in mT/metre. The slew rate affects how fast the gradient can be turned on, measured in mT/m/s, how fast in general things get done (influences the minimum TE and TR) limits the number of pulse sequences that may be played out. In turn, the amplitude mainly affects the minimum slice thickness and FOV: the larger the gradient, the smaller the FOV and the smaller the slice thickness.

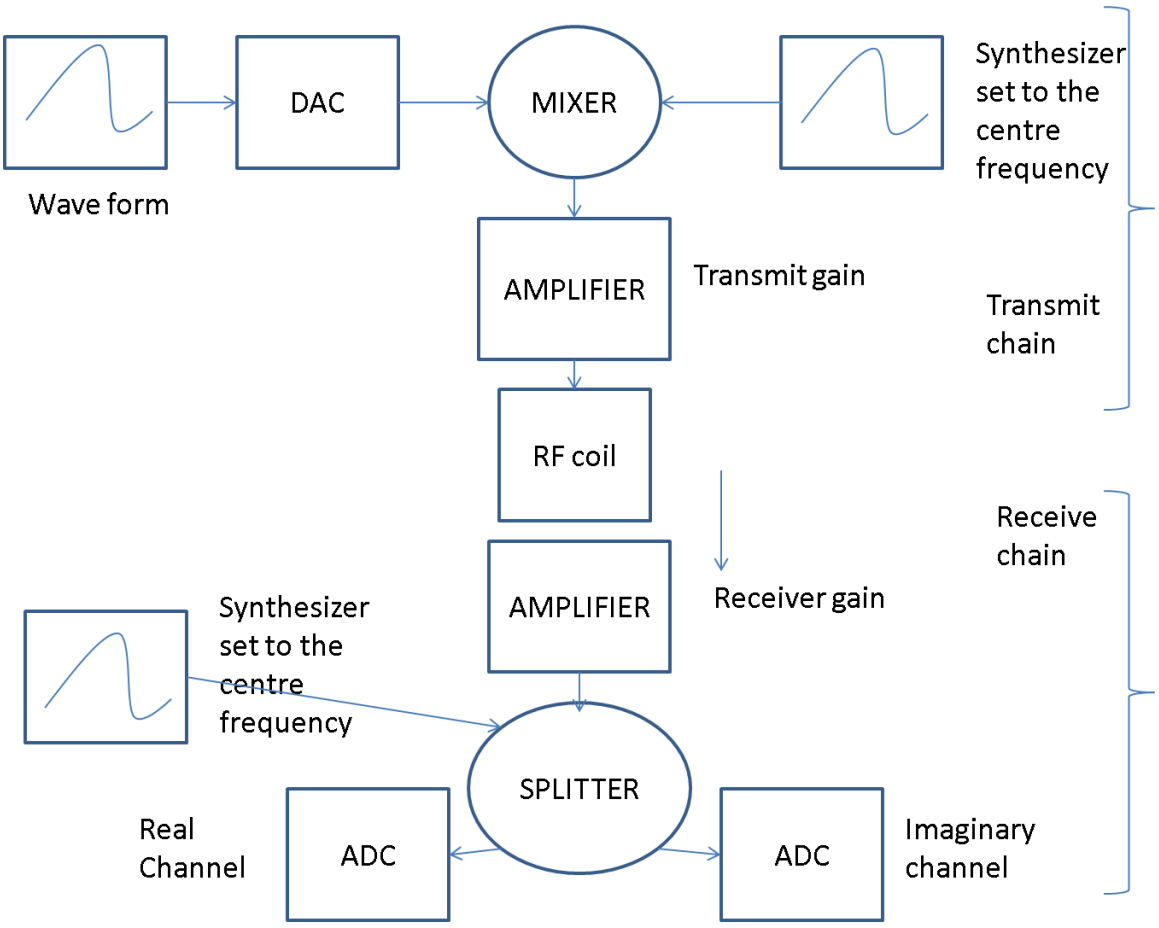


Figure 2.8: The RF system of an MRI system. The image was re-drawn after a figure in (NessAiver, 1997).

There are not only technical limitation to the slew rate, they effectively could also cause peripheral nerve stimulation (PNS).

The RF system is one of the most important parts of the system. Figure 2.8 diagrams the complex RF generation process, whose details are beyond the scope of this thesis. A fuller discussion of RF coils includes their shape, polarisation, and number of separate elements. The shapes of RF coils include: saddle coils, surface coils, linear RF coils, and quadrature coils. Saddle coils have a uniform field strength within a volume, so they are also called volume coils. Surface coils (they will come up later) give high signal to noise ratio (SNR) close to the coil, but their sensitivity fades away quickly. They have been used almost exclusively as receive coils, but since

the early twenty first century, when the parallel transmit concept appeared, they have been re-considered as transmit elements as well. The rule of thumb is that their sensitivity is only one diameter from the centre of the coil. Nevertheless, multiple elements can be placed next to each other, providing high resolution imaging with the limitations as mentioned above (non-uniform sensitivity, very high sensitivity next to the coil and very fast drop off after). Surface coils are used for liver imaging because the liver provides easy access to them and in turn the surface coils provide much better SNR than the body coils for receive only.

Finally, the list is not complete without the MR bias field. This is a result of small inhomogeneities in the magnetic field that causes slow variations of intensity in the image domain, which in general is modeled with a polynomial. This effect often hinders very important features for example in case of the spine, liver and breast MR imaging.

Linearly polarised surface coils instead are sensitive to magnetic fields along one axis, while quadrature coils are sensitive along two orthogonal direction. They are circularly polarized and always appear in pairs. One channel is aligned with the reference signal and one channel is offset with 90 degrees to the reference signal, it has a 90 degree phase shift. Phased array coils give very high SNR, have better properties than simple surface coils, and that is why they are used in spinal imaging and during abdominal scan which is almost always degraded by patient and bowel movements.

2.3 MRI of the liver

Texture is important for interpretation of clinical liver images, since it has spatial organisation that shows both structural and statistical characteristics (for example, see Figure 2.9). As a result, the larger the area that is available for analysis, the

better to evaluate certain image analysis methods.

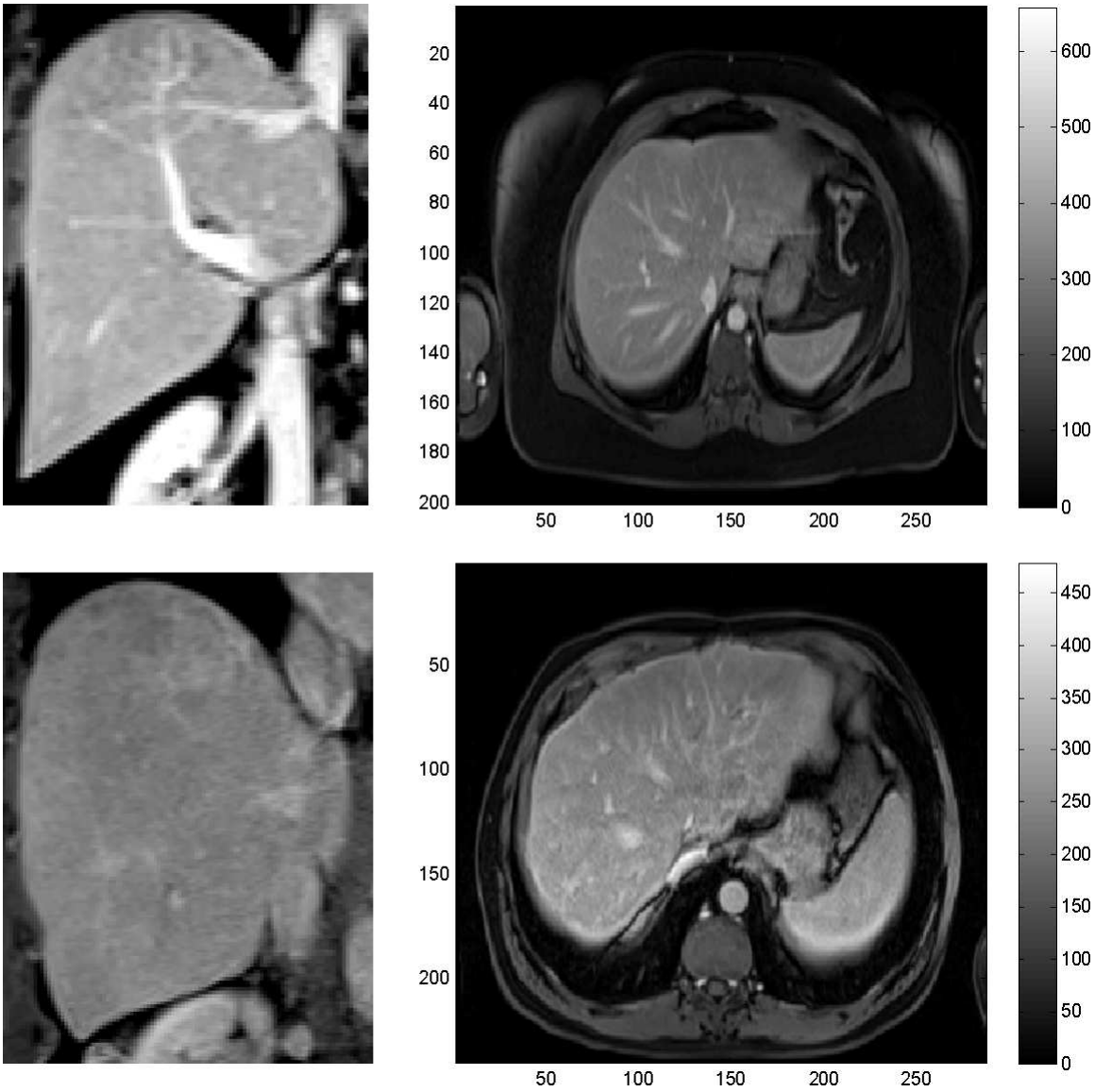


Figure 2.9: Human liver MRI at 1.5T, coronal and axial view. The images were kindly provided by Acuitas Medical, Swansea and acquired at the University of Arizona in 2005. The top line shows a healthy liver, while the second line a Grade 4 liver disease.

2.3.1 Why is the liver a good organ for studying structural changes?

The liver is a relatively homogeneous organ (shown in Figure 2.10), so it is ideal for the detection of texture changes. In contrast to the pancreas, it does not itself contain

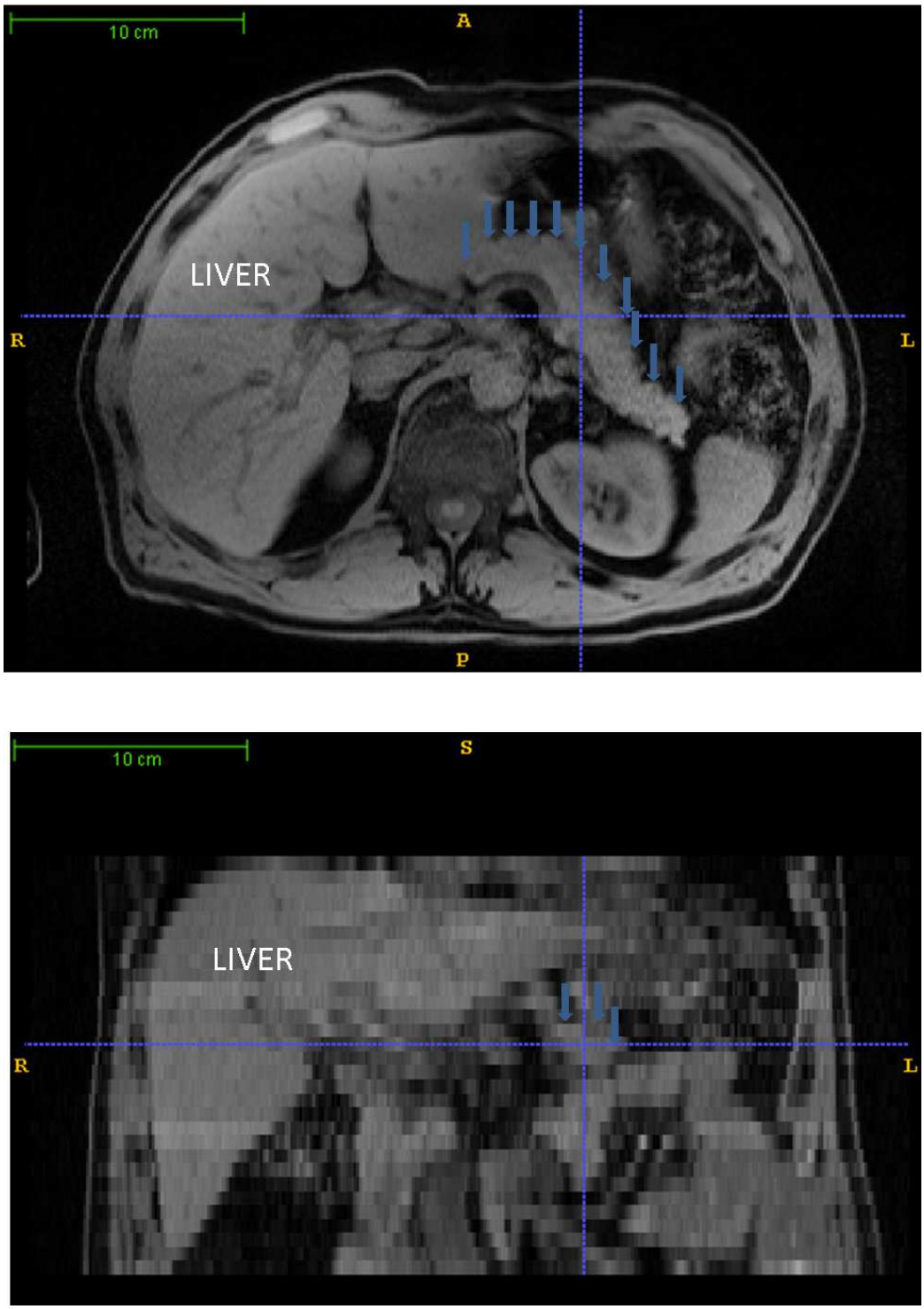


Figure 2.10: Liver and pancreas appearance on MR images: axial and coronal view. The figure is annotated for the liver (in white) and the pancreas on both images shown by the blue pointer; the upper boundary of the pancreas is marked by the blue arrows. The images were acquired at the Churchill Hospital, Oxford, UK.

very strong structural changes in its natural form. Nor is it hard even to identify it on MR images. There are several papers that suggest that the biology underlying the formation of fibrosis, with the help of the stellate cells as a result of injury, is very similar both in the liver and in the pancreas, and that they follow the same cellular pathway (Jaster et al., 2002) (McCarrol et al., 2004).

2.3.2 Specific coils that are used

Since the liver is in the abdomen, MR imaging inevitably suffers from two types of motion: the first is patient motion; the second is natural bowel movements, notwithstanding the respiratory effects. For these reasons, liver scans use specialised surface coil arrays which are placed on the abdomen of the patient. As noted earlier, surface coils have been used for our liver MRI study to maximize the SNR.

2.3.3 Inflammation and fibrosis imaging

Inflammation and fibrosis imaging is generally considered to be easier with PET (Signore et al., 2000) (Signore et al., 2010) than MRI (Jones, 2005) (Gareau et al., 2001). However, as we will show later, MRI detects a set of features and, textural changes during disease progression that are quantifiable even without specific markers. Another view is that fibrosis is nothing other than scarring, as a result of which both the density increases and the water content decreases. From this perspective, all imaging modalities should to a certain extent show its presence.

2.4 Conclusion

In this Chapter we have described the workings of an MRI system: both the hardware and the methodologies used to obtain images from an MRI system. We also sketched two possible pulse sequences that may be used to image inflammation/fibrosis imaging

with MRI.

We conclude by noting that although MRI has a very wide range of applications, it might eventually prove not be the best modality for our clinical purposes. Nevertheless, it is currently the best option at present. It provides the necessary information that is required for our studies, taking also in account the patient's perspective (using a non-invasive imaging modality).

Chapter 3

Phase and structure identification

The notion of phase appears in all areas of natural sciences. It also has a specific role in image representation and coding. We differentiate between absolute and relative phase, in effect phase means some type of relative measure with respect a reference signal. Given an arbitrary n-D signal the so called "structure" of the signal, or the ensemble of interesting events, are encoded by the spectral phase. In this chapter after a short introduction we focus on the local phase of signals and we complete the chapter with the presentation of two isotropic model-free feature detectors.

3.1 From global to local analysis

The Fourier Transform (FT) allows the decomposition of a signal with the basis of infinite complex oscillating trigonometric functions:

$$F[f](u) = \int_{-\infty}^{+\infty} f(t) \cdot e^{-jut} dt \tag{3.1}$$

where the lower case f denotes the time domain function and F is its FT, j is the complex imaginary unit. This model allows to derive global signal shape characteristics. It uses the *cos* function to probe the signal for even components (i.e. functions

which satisfy $f(x) = f(-x)$), and an imaginary *sin* function to quantify the odd components ($f(x) = -f(-x)$).

Information content of the images are retained in the phase of FT spectra as opposed to its magnitude. This fact was demonstrated in (Oppenheim and Lim, 1981) and is presented here in Figure 3.1. The two images were reconstructed with interchanging the phase information. As it suggests the phase information dictates which organ dominates the field of view. However, the FT only allows the derivation of the global phase of the images, which is a major drawback for many image analysis applications.

In particular, medical image analysis relies on the accurate identification of localized features of interest. Features can be organ boundaries or disease bio-markers within the tissues, for example. Applications that are of great importance within this context is organ or tumour segmentation, registration or tissue characterisation based on the extracted markers. This latter aspect is of primary interest to us.

Local feature extraction is possible with sampling and analysing only a small portion of the signal at a time. The sampling window can range from a rectangle to a Gaussian function and the transform is then defined as:

$$W[f](t_0, u_0) = \int_{-\infty}^{+\infty} f(t) \cdot g(t - t_0) e^{j u_0 t} dt \tag{3.2}$$

This is also called windowed FT (WFT) or Short Time FT (STFT). The Gaussian as a windowing function is a smart choice, because this is considered to be the only kernel that has an optimal spatial (temporal) and spectral domain localisation at the same time. The Heisenberg uncertainty gives the lower bound on the obtainable signal localisation $\sigma_t^2 \sigma_u^2 \geq \frac{1}{4}$. This localisation now allows the derivation of a local phase and local energy of the signal. Unfortunately, this construct is still limited because the sampling window has a fixed size and general image features have their

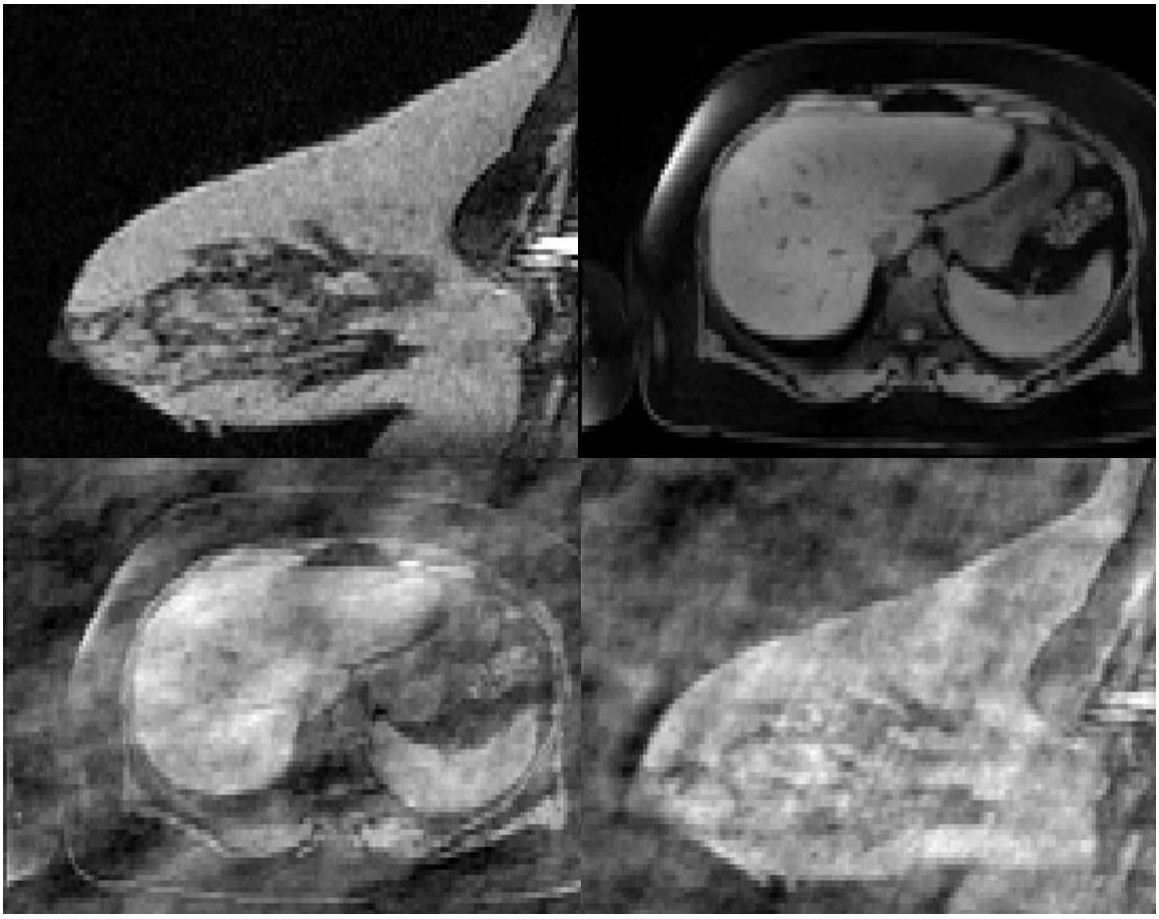


Figure 3.1: The importance of Fourier phase. The first row displays the original MRI images of a breast and liver. The second row in order shows when the Fourier phases of the two images are interchanged.

own characteristic scale (Lindeberg, 1996).

Wavelets allow the analysis of the signal/image using variable size windows using a scaling function and a mother wavelet. The correlation between the signal and the wavelet gives rise to the measurement values. The general rule with time(space)-frequency analysis is that the more we localize in time the less certain we are in the identified spectra and vice versa. The definition of the FT earlier shows that a good frequency identification would ideally require an infinite sample size, which means that there is no temporal(spatial) localisation whatsoever. The reverse follows the same argument, the more we localise in time(space) the less certain we are about the frequency content of the signal.

In terms of filtering operations the wavelet decomposition is achieved with a bank of low and high-pass filters. These yield a bandpass representation of the underlying n-D signals. A mathematically less rigorous, but in medical image analysis even greater importance has the local feature analysis based on specific bandpass filter decompositions that match the desired feature signal shapes.

The derivation of local image descriptors has two main steps:

1. Decompose the image using an appropriate bandpass filter.

2. Use a signal model that allows the derivation of the respective image features.

In what follows, after a short introduction, we review the state-of-the-art local signal models that allow the definition of local phase and the phase congruence based feature model.

3.2 Local phase based structural description

General phase based signal processing seeks to find the linear phase $\phi \in [0, 2\pi]$ and local amplitude $a \in \mathbb{R}$ of a signal $f(x) = a\cos(kx + \phi)$, with $x \in \mathbb{R}$ and $k \in \mathbb{R}_+$.

The importance of local phase information in the description of signal structures has been long recognized in (Gabor, 1946) in the context of microscopy images, however, it was not until the seminal communication of (Oppenheim and Lim, 1981) that the concept was re-discovered for image analysis.

In (Oppenheim and Lim, 1981) it was demonstrated that signals may be accurately reconstructed up to a constant factor using their phase information and unit magnitude information. The criteria under which the Fourier phase allows this is bound to the finite length of the sequence; which applies to 1-D and m-D signals (Hayes, 1981). The role of the Fourier magnitude was shown to be a surrogate for signal reconstruction under specific conditions. However, in general, the roles of phase and magnitude cannot be interchanged as we demonstrated in Figure 3.1.

The same concept, based on the signal's zero-crossing, was developed in the field of communication theory (Curtis, 1985). Requirements under which a signal's zero crossings, first order derivative and its relation to its Hilbert transform allows optimal reconstruction leads to Logan's theorem (Fitzgerald, 1991), which states that any bandpass signal can be reconstructed up to a constant factor by its zero crossings if the bandwidth of the signal is less than one octave. The validity and extension of this 1-D constraint became one of the central themes in the examination of Marr's zero crossings and primal sketch in 2-D (Marr, 1982). Related to this is the Canny edge detector illustrated in Figure 3.2. This figure shows that the boundaries and the

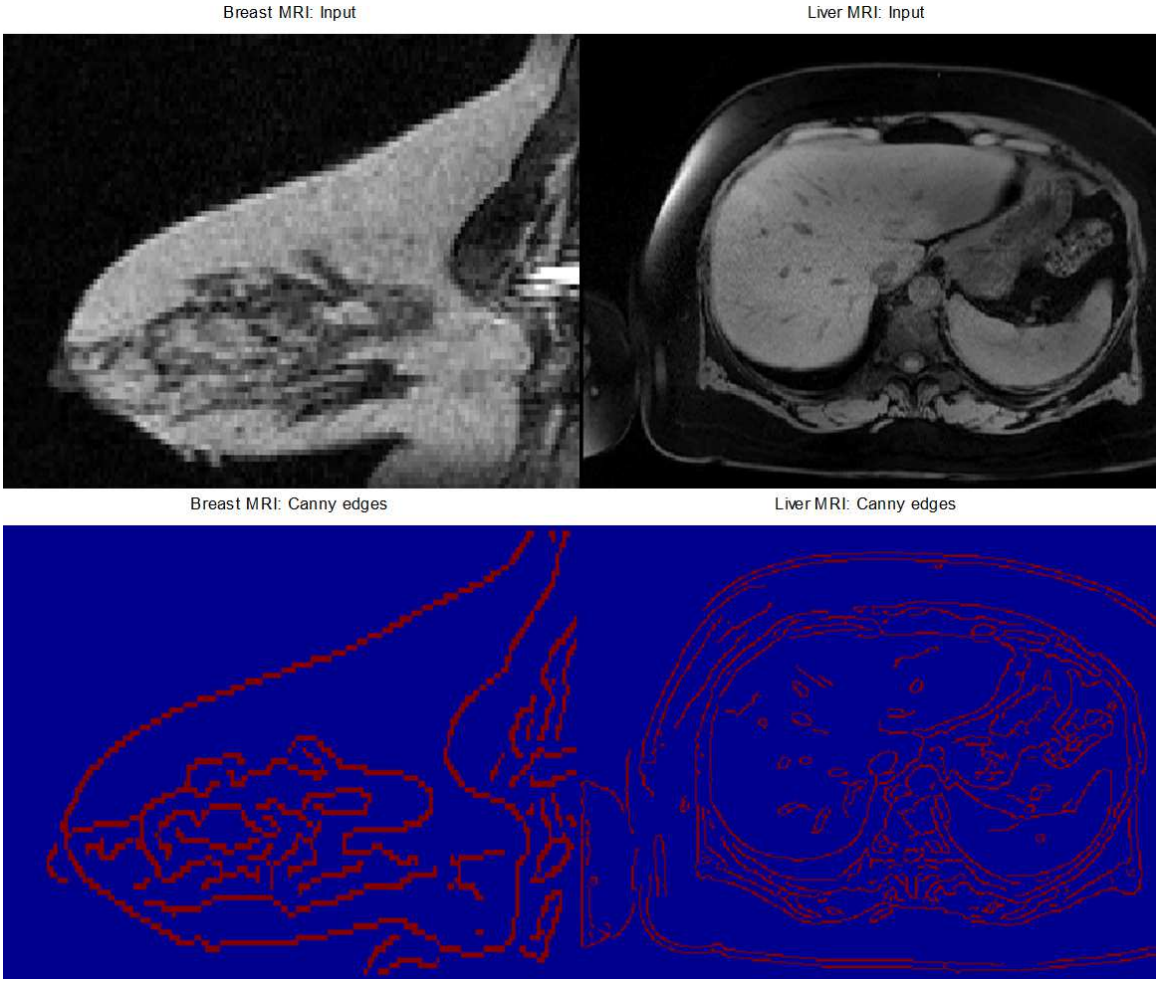


Figure 3.2: The Canny edge detector (Canny, 1986) applied to breast and liver MRI. The top row shows the input measurements, while the second row is the associated edge map.

major structures of the respective MRI breast and liver images can be obtained.

The above investigations led to a separate branch of image processing that focuses on phase and zero crossings based image reconstruction (Hurt, 1989). As opposed to this higher level image processing task, we take a step back and examine the use of phase in the low-level characterization of 2-D signals.

For optimal feature detection one seeks a decomposition that is orthogonal and spans the signal's space. One way to achieve this is by the use of a pair of quadrature filters that allows the definition of phase and signal energy while satisfying the invariance-equivariance property. Equivariance means that the phase value changes monotonically with the structure type, as does the energy with feature certainty and local orientation with the geometric orientation. However, this change is by no means dependent on the actual phase representation.

Phase information carries parity symmetry information about the underlying structure, Figure 3.3, i.e. it gives information regarding its oddness/evenness, while signal energy gives a measure of signal certainty.

Figure 3.3 shows how we could relate the quadrature filter output measurements to the phase and signal shape. As it will be presented later the local phase is derived as the angle between the even and odd filter outputs. As an analogy if the input is a pure odd function then we have a high value for the respective filter and close to zero on the quadrature pair. This means that we evaluate the *arctan* function at infinity, which is $\frac{(2k+1)\pi}{2}$ with $k \in \mathbb{Z}$. This is why we associate an odd signal shape with a $\frac{\pi}{2}$ and $\frac{3\pi}{2}$ phase value.

Given a locally 1D signal model $f(x) = a\cos(kx+\phi)$, $x \in \mathbb{R}$, $k \in \mathbb{R}_+$, the analytic signal is defined as:

$$f_A(x) = a[\cos(kx+\phi) + i\sin(kx+\phi)] \tag{3.3}$$

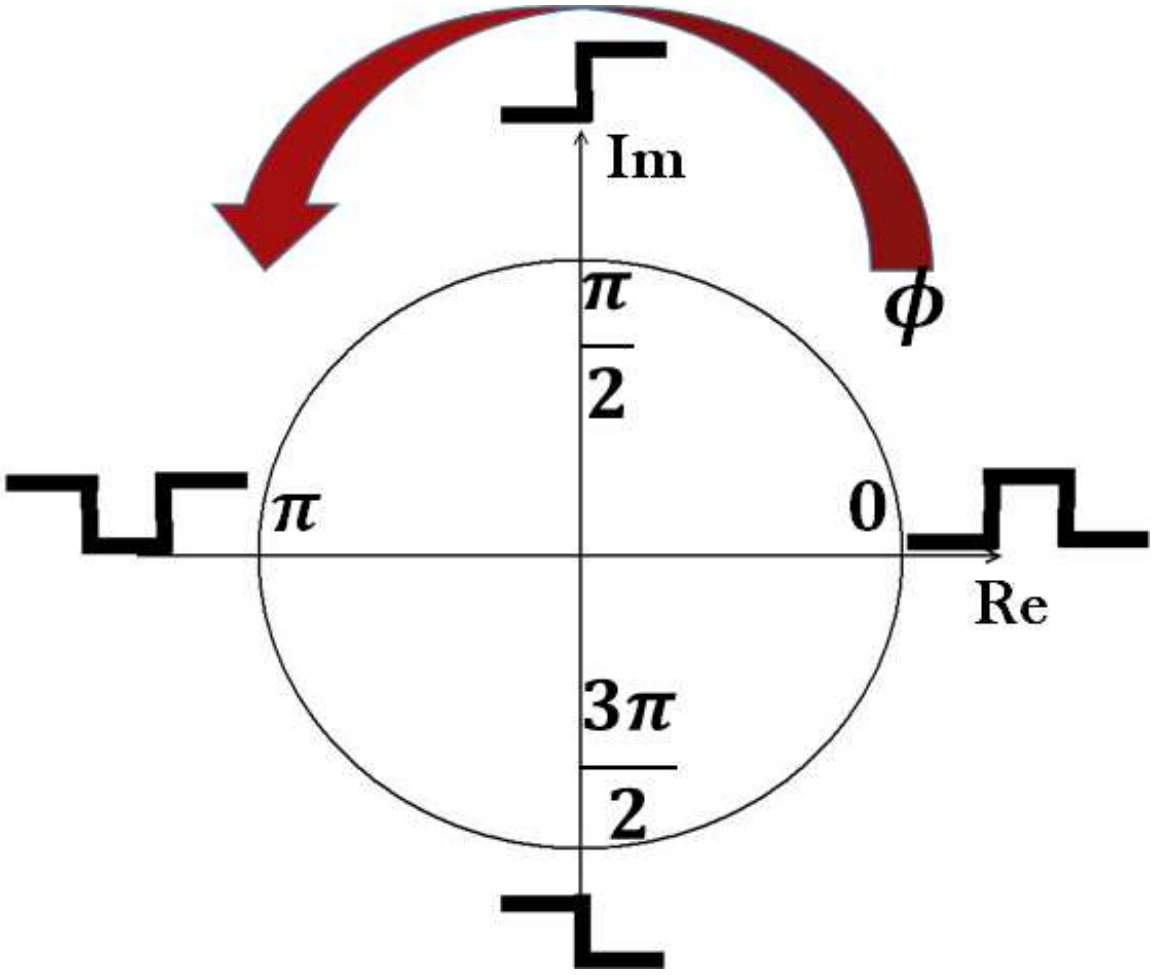


Figure 3.3: Local phase and signal parity. The major part of this representation was first proposed in (Granlund and Knutsson, 1995).

where the even and odd components are in quadrature: they carry the same frequency content, however their phase is shifted by $\frac{\pi}{2}$. For general signals, the odd component is achieved by the use of the Hilbert transform. In the Fourier domain the Hilbert transform is defined as the transfer function (Bracewell, 2000):

$$H(u) = -i \text{ sign}(u) \tag{3.4}$$

where u denotes frequency and the operator is an all-pass operator that suppresses the DC component and, most importantly, it leads to a one sided spectrum of the analytic signal, i.e. $F_A(u) = 0$ for $u < 0$.

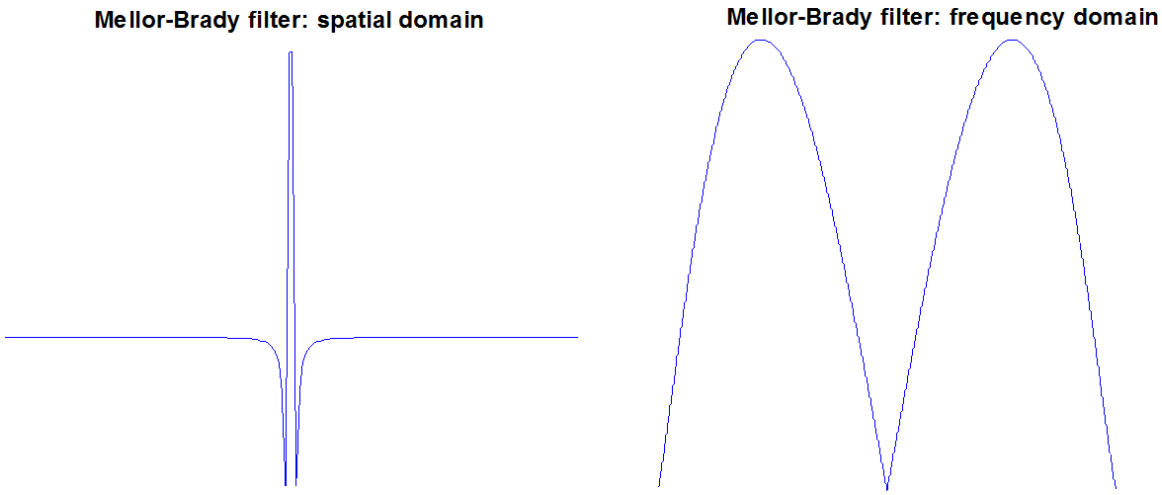


Figure 3.4: The Mellor-Brady filter in the spatial and Fourier domain as defined in the text.

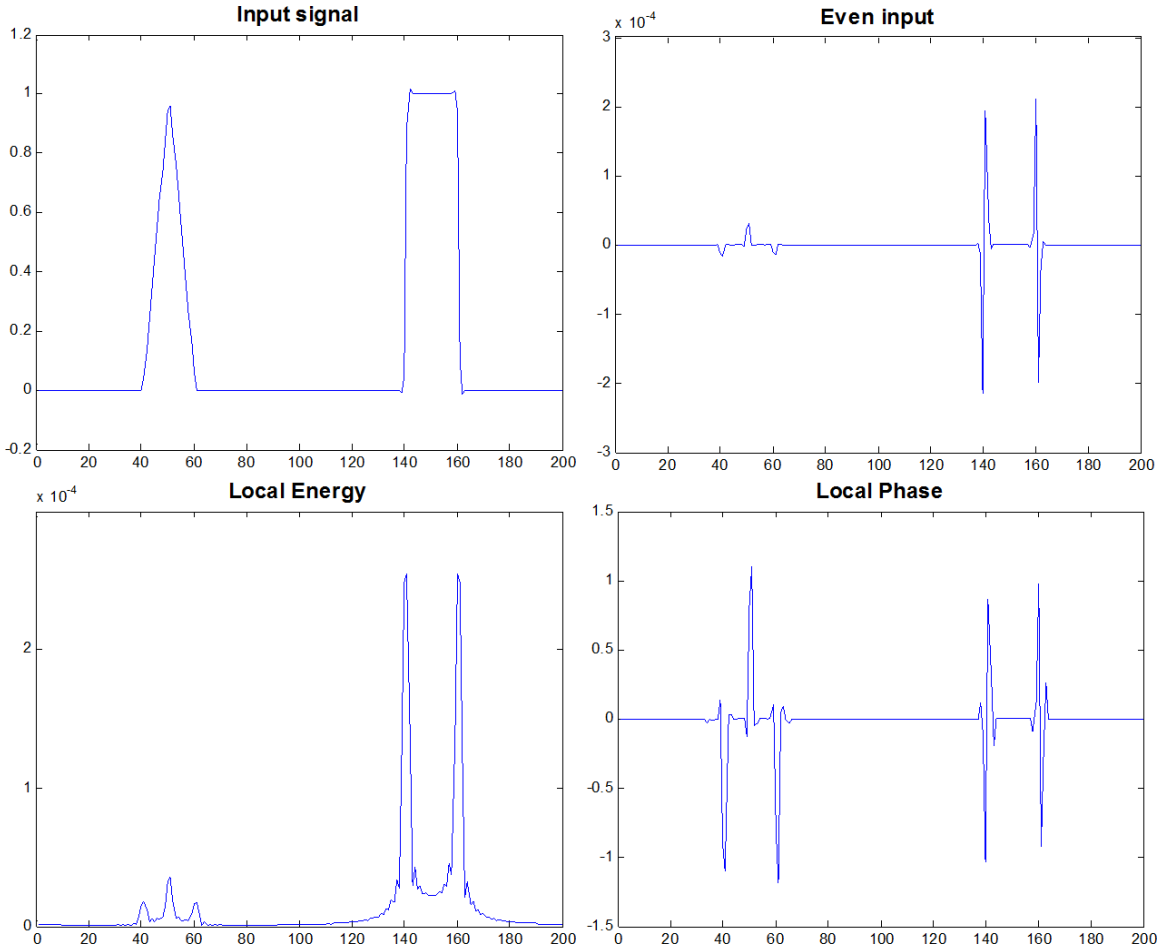


Figure 3.5: Illustration of the analytic signal for a 1D signal with even and odd components.

Based on the above, the 1D analytic signal with the instantaneous phase and energy is derived as:

$$\begin{aligned} \phi(x) &= \text{atan}\left(\frac{\mathcal{H}[f](x)}{f(x)}\right) \\ a(x) &= \sqrt{f(x)^2 + \mathcal{H}[f](x)^2} \end{aligned} \tag{3.5}$$

with $\mathcal{H}$ denoting the Hilbert transform.

The bandpass filter that we pre-dominantly used in this thesis in the Mellor-Brady filter. The spatial and frequency characteristics of this filter are presented in Figure 3.4. This shows that the spatial domain filter shape resembles that of a Difference of Gaussian (DoG), but with a much better spatial localisation. This property however is counter balanced with the broad spectrum in the frequency domain.

Example analytic signal features for the ideal triangular and rectangle waveform are summarised in Figure 3.5. This shows that the most impartial feature representation is achieved with the instantaneous phase as compared to the instantaneous energy.

The 2-D extension of the above model necessitated two developments. First, extension of the Hilbert transform in the form of the Riesz transform (Stein, 1970). The second, an appropriate signal model formulated using quaternions (Felsberg and Sommer, 2002)(Perwass, 2009) and potential field theory which in 2-D yields the monogenic signal. The study of potential theory led to an alternative extension of the Hilbert transform into higher dimensions in the Fourier domain with the realization that higher dimensional Hilbert transforms are possible with the definition of the sign function in higher dimensions, Figure 3.6. The Riesz transform in the Fourier domain is defined as:

$$\mathbf{R}(u_1, u_2) = \left(\frac{iu_1}{\sqrt{u_1^2 + u_2^2}}, \frac{ju_2}{\sqrt{u_1^2 + u_2^2}}\right) \tag{3.6}$$

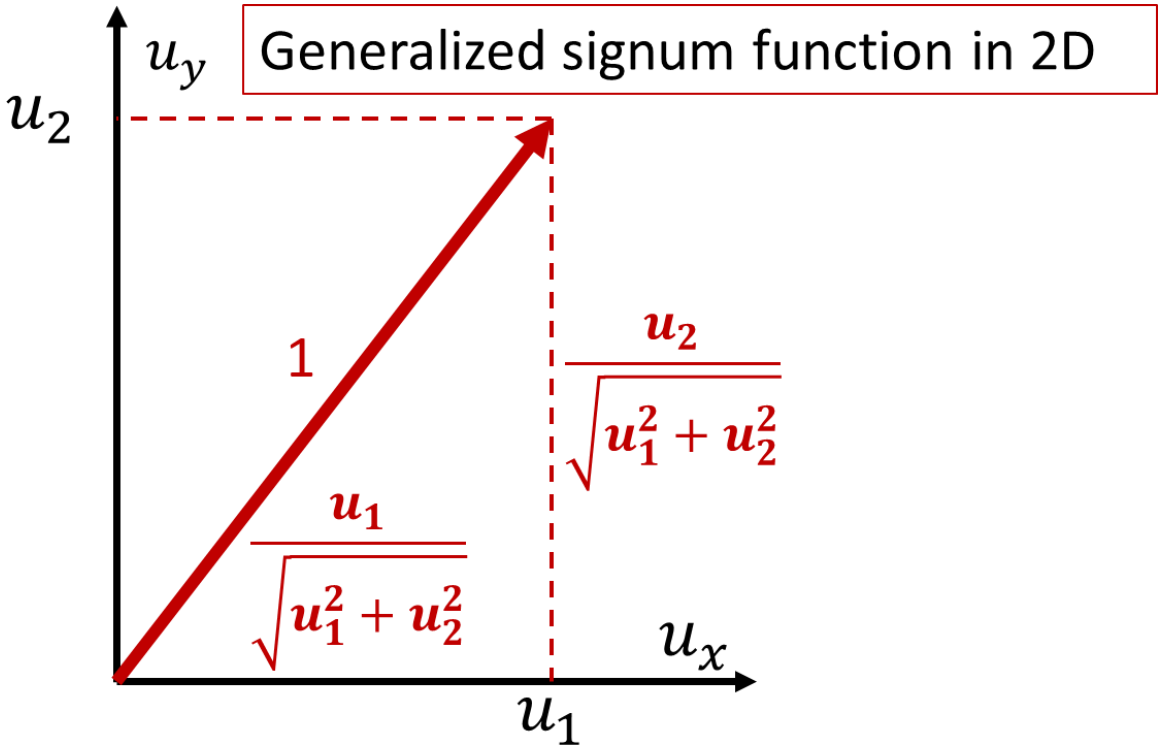


Figure 3.6: The signum function generalized to 2D. The figure was re-drawn after the image presented in (Nabighian, 1984).

where i and j are quaternion units, and u_1, u_2 denote frequencies.

Given that we wish to perform local signal analysis the images are first filtered to select a range of frequencies and localize the spatial signal using a bandpass filter. This could be a Log-Gabor, Cauchy, DoG, Mellor-Brady or any other bandpass filters.

3.2.1 2-D Monogenic Signal

The first fully isotropic signal model to make use of the Riesz transform is the monogenic signal. This assumes an underlying signal model:

$$f(x, y) = a\ \cos(x\ \cos(\theta) + y\ \sin(\theta) + \phi) \tag{3.7}$$

where a is the amplitude of signal, x, y are 2D Cartesian coordinates, θ is the local orientation of the signal and ϕ denotes phase. Note that this model assumes that a

single orientation is present.

The monogenic signal (Felsberg and Sommer, 2002) is derived by the following steps. First the bandpass signal is defined as $b\left(f(\overline{x})\right)$, where $\overline{x} = (x, y)$ denotes the 2-D image coordinates. Following this, the Riesz responses are calculated as:

$$(r_1, r_2) = (R_1(b\left(f(\overline{x})\right)), R_2(b\left(f(\overline{x})\right))) \tag{3.8}$$

where r_1, r_2 denote the Riesz response in the spatial domain and R_1, R_2 are the Riesz filters. The new signal model lives on a sphere, see Figure 3.7, parameterized by its radius and two angles.

The monogenic signal is then defined as the three-vector in the spatial domain:

$$m(b(f)) = (b(f), r_1(b(f)), r_2(b(f))) \tag{3.9}$$

where m denotes the monogenic signal, and for ease of notation we dropped the spatial coordinates. The respective calculations in the spatial domain is presented by the diagram in Figure 3.8. The spatial domain 2-D filter representations are illustrated in Figure 3.9, where the two Riesz filters are a sophisticated counterpart of 2-D differentiators. In this figure white means positive and black means negative value. The FT domain representation of the Riesz filters without the bandpass effect is represented in Figure 3.10.

Based on this filter triplet (Figure 3.9 and 3.10) the monogenic components are derived as:

$$\begin{aligned} A(b(f)) &= \sqrt{b(f)^2 + r_1(b(f))^2 + r_2(b(f))^2} \\ \theta(b(f)) &= \text{atan2}\left(\frac{r_2(b(f))}{r_1(b(f))}\right) \\ \phi(b(f)) &= \text{atan}\left(\frac{b(f)}{\sqrt{r_1(b(f))^2 + r_2(b(f))^2}}\right) \end{aligned} \tag{3.10}$$

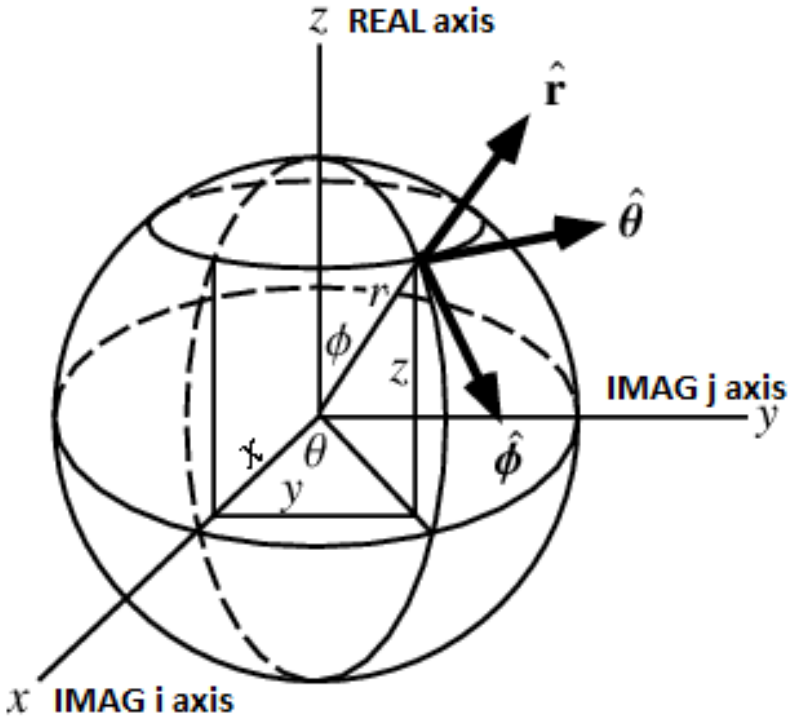


Figure 3.7: The monogenic signal as was originally proposed in (Felsberg and Sommer, 2002). This image was adapted without permission from (Mathworld, 2014).

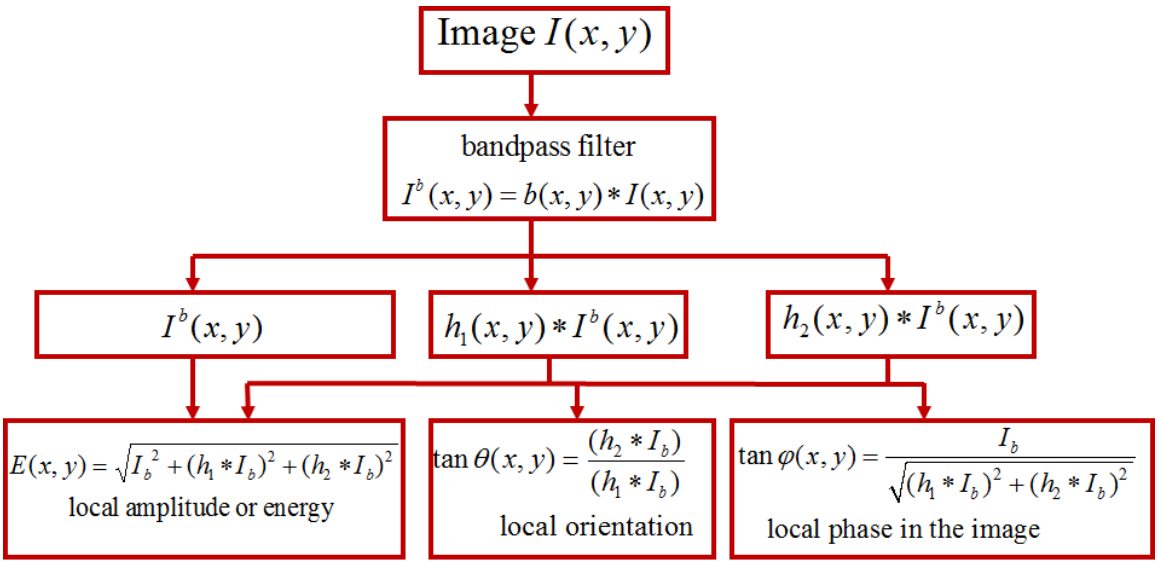


Figure 3.8: Monogenic computation in the spatial domain. This image was provided by Professor Sir J. Michael Brady.

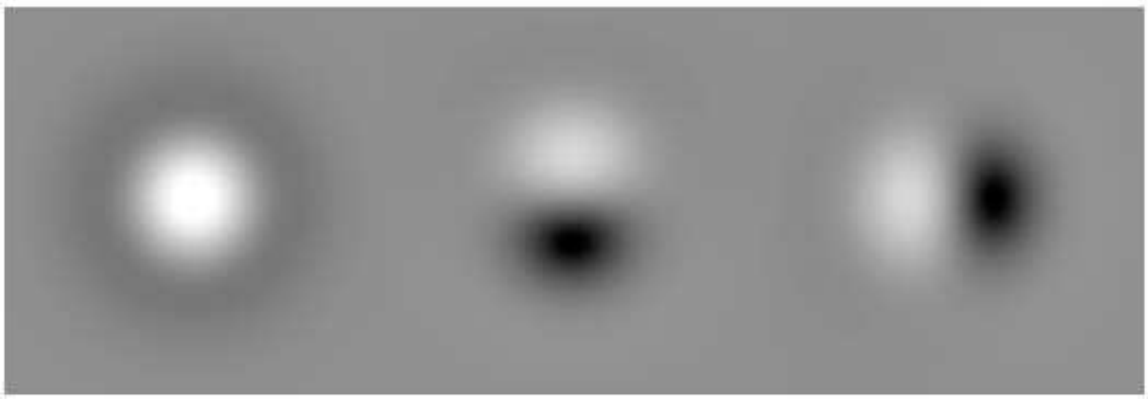


Figure 3.9: Monogenic triple filters. From left: bandpass filter as a Difference of Gaussians (DoG), $\frac{x}{2\pi\sqrt{(x^2+y^2)^3}} * \text{DoG}$ (horizontal orientation) and $\frac{y}{2\pi\sqrt{(x^2+y^2)^3}} * \text{DoG}$ (vertical orientation). This image was provided by Professor Sir J. Michael Brady.

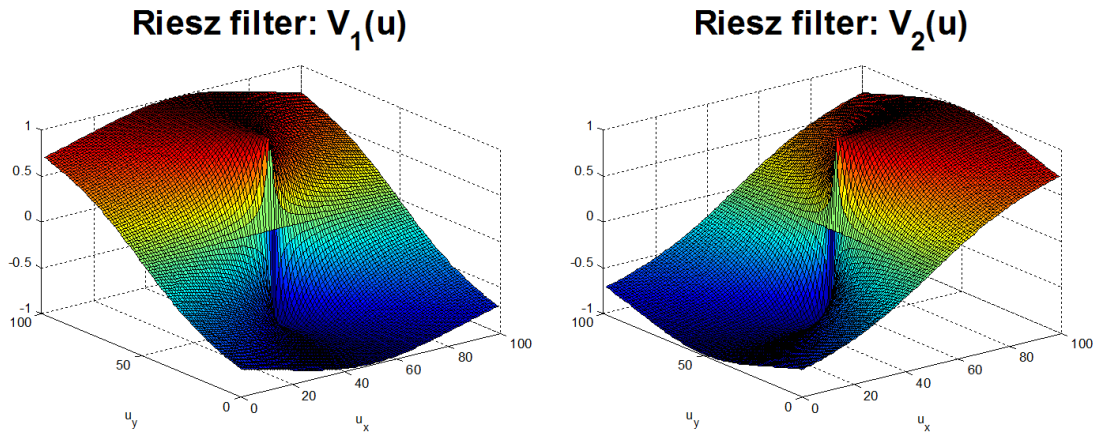


Figure 3.10: Riesz filters in 2D in the frequency domain.

where A is local energy, θ denotes the local orientation, ϕ is the local phase, atan is defined on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and atan2 maps down values to the interval $[-\pi, \pi]$. These satisfy the principle of split-of-identity (Felsberg and Sommer, 2002). This term means that by examining a narrow band signal we have an orthogonal decomposition into local energy, phase and orientation. In effect, these descriptors are characterized by the invariance-equivariance property, which means, for example, that the energy change is invariant to structural changes described by the phase and it only detects the signal certainty. In a similar manner, the phase only changes with the structure as opposed to any other signal quality.

MONOGENIC SIGNAL computation in 2D in the Fourier domain using MATLAB (OCTAVE) syntax	
1. Define the mesh parameterized by the 2-D Fourier variables	[u1, u2] = meshgrid(linspace(-maxu, maxu, dimy*2), linspace(-maxu, maxu,dimx*2)); hyp = sqrt(u1.^2+u2.^2);
2.Define the bandpass filter (kern in the spatial domain) and the Riesz filters in the FT space (V1 and V2)	V1 = i*u1./(hyp+eps); V2 = i*u2./(hyp+eps);
3.Derive the bandpass image: pointwise multiplication	im_r = real(IFFT2(fftimg.*FFT2(ifftshift(kern))));
4.Derive the Riesz components of the image	r1 = real(IFFT2(ifftshift(V1.*fftimg))); r2 = real(IFFT2(ifftshift(V2.*fftimg)));
5.Derive the monogenic components	See text for details.

Figure 3.11: Step by step computation of the monogenic signal in the Fourier domain. For details, see text.

Due to the rotation and linear shift invariance of the generalized Hilbert transform in Euclidean space, the m-D analytic signal features are also rotation and translation invariant.

The monogenic signal computation is implemented in the Fourier domain by the steps presented in Table 3.11. In step 2 the variable *kern* denotes the spatial domain definition of the selected bandpass filter. Throughout the computations we have used the *Mellor-Brady* filter unless otherwise stated. In step 3 the bandpass filtered image is derived. The variable *fftimg* denotes the 2-D FT of this function as it appears in step 4. The descriptors are derived following the equations presented in eq. 3.10 with the specific details given below the equations.

The monogenic features are illustrated for a healthy and a diseased liver image in Figure 3.12. These set of monogenic descriptors suggest that differences between the healthy and diseased liver stage can be detected. In particular, while the healthy liver tissue shows some *regularity* in its architecture along with clear blood vessels and ducts, the diseased liver parenchyma becomes more homogeneous. This is against

the philosophical assumption that disease progression *always* is associated with a transition from *order* to *disorder*. The resolution of the images also influence this, and in this particular scenario the tissue appears to move towards a more *homogeneous* representation by the gradual break-down of the healthy liver architecture.

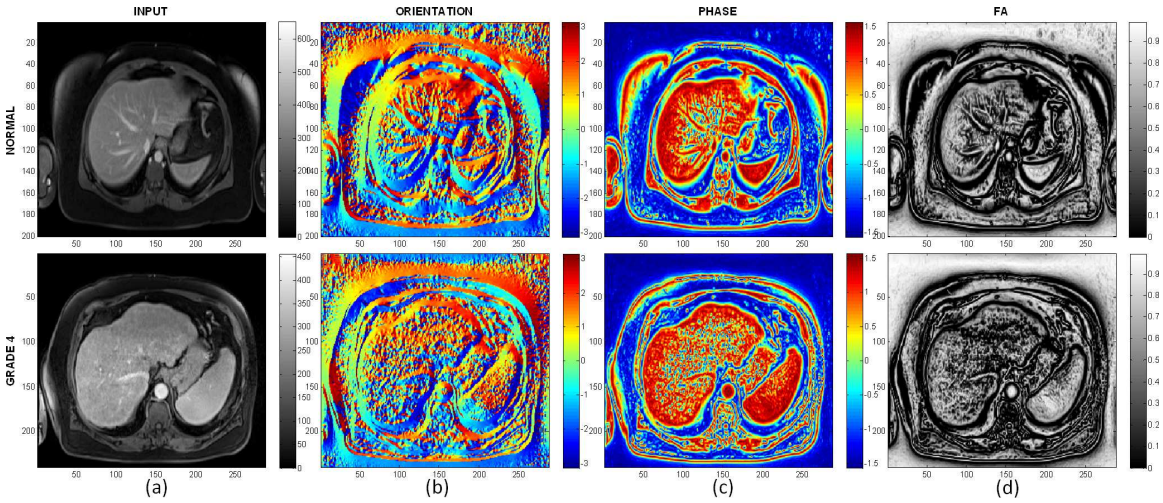


Figure 3.12: 2D Monogenic signal derived features for a healthy and a Grade 4 liver disease patient, from 1.5T MRI images. The intensity images are shown in (a), the local orientation (b), the local phase maps on (c), and the feature asymmetry maps derived from four consecutive scales on (d). Notice the differences in the structure of the liver parenchyma, especially on the FA and phase maps.

3.2.2 2-D Analytic Signal Geometry

The Riesz transform and the above signal model provided the possibility for the isotropic extension of the 1-D analytic signal to m-D. However, given that the Riesz transform is a function of the Radon transform (which relies on 1-D line integrals) the above model is limited to the analysis of i1D features (Felsberg and Sommer, 2002)[eq. 20]:

$$\mathcal{R}\left\{f_{\mathcal{R}}\right\}(t,\theta) = (i,j)n_\theta h_1(t) \star \mathcal{R}\left\{f\right\}(t,\theta) \tag{3.11}$$

where $\mathcal{R}$ denotes the Radon transform, f is the 2-D image, $f_{\mathcal{R}}$ is the 2-D Riesz transform of the image, n_θ is the normal, θ orientation and h_1 is the 1-D Hilbert

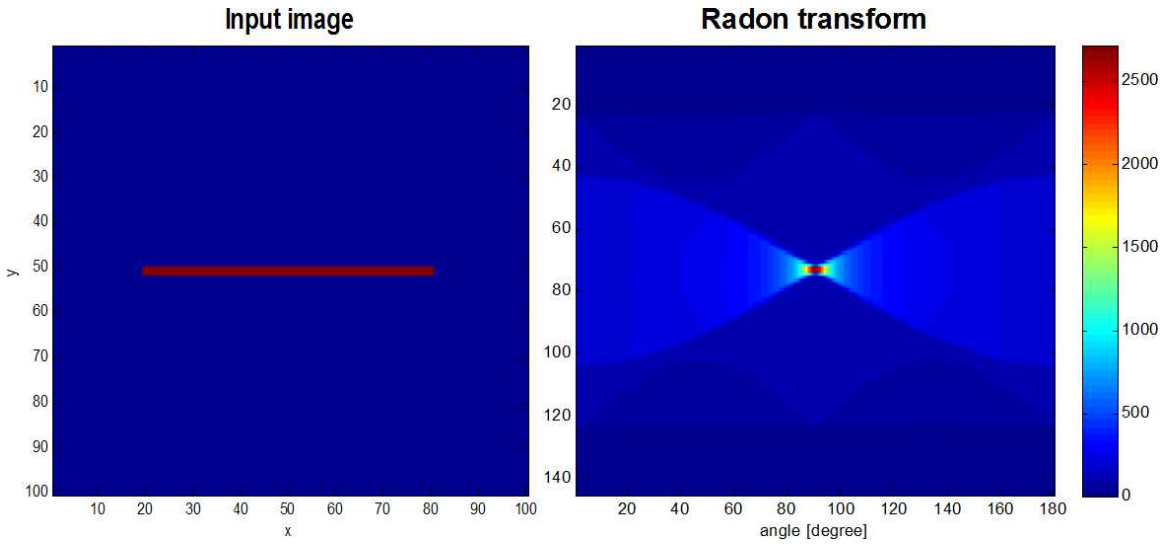


Figure 3.13: i1D line profile in 1-D and the corresponding Radon transform. Note that the Radon transform has a particularly broad spectrum.

transform. For proof see the above reference.

This means that the Radon transform of the Riesz transform of the image can be calculated with the 1-D Hilbert transform applied to the Radon transform of the image. As a result, there is no possibility to describe the relationship between two intersecting i1D lines, a typical i2D feature that would be expected to be detected by "corner" detectors.

The realization that 2-D signals exist in a 3-D projective subspace of the homogeneous conformal space provides a new perspective on the geometry of the 2-D signals (Wietzke et al., 2009). Claim supported by the derivations in (Wietzke and Sommer, 2008). The earlier i1D signal model (eq. 3.7) is extended as the superposition of i1D features, which now reads as:

$$f(x,y) = a \sum_{i}^{2} \cos(x\ \cos(\theta_i) + y\ \sin(\theta_i) + \phi) \tag{3.12}$$

with the same notation as in eq.3.7.

Within this model any i2D may be represented as the sum of i1D signals. In fact, the 2-D analytic signal naturally includes the framework of the ciD and no knowledge

of the iD property is needed prior to the computation itself. However, this proposed extension of the monogenic signal has its own limitations. For finite signals, for example, the apex angle is never exactly zero for an ideal i1D, because the i1D signal in Radon space has a broad spectrum (see Figure 3.13) that is picked up by the 2-D analytic signal, Figure 3.14. The apex angle is the smallest angle between two linear structures whose exact calculation is defined later in this chapter. In contrast, the monogenic signal assumes a locally 1-D signal, therefore it expects a zero apex angle at each location in the image. The apex angle calculated by the 2-D analytic signal is in fact influenced by the width of the signal that is analyzed and changes with scale, which is not necessarily a desirable feature property.

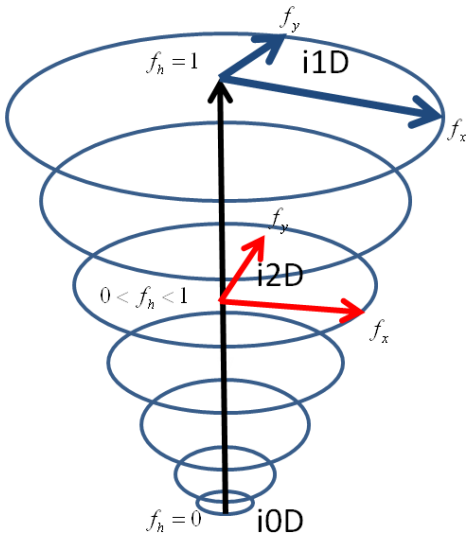


Figure 3.14: 2D analytic signal. Extension with the homogeneous component, which now includes all three classes of ideal and non-ideal structural features i0D, i1D and i2D as indicated on the figure. The bottom part of the spiral includes the i0D, the top plane includes the i1D, while the i2D features span the whole spiral as is indicated by the labels. f_x, f_y are the conjugate Poisson components. Figure re-drawn after (Wietzke et al., 2009)

In 2D Radon space the resulting system of equations for the conjugate signal components are represented as:

$$\begin{aligned} f_x &= \sum_{i=1}^{n} \mathrm{a} \cdot \cos(\theta_i) \cdot \sin(\phi) \\ f_y &= \sum_{i=1}^{n} \mathrm{a} \cdot \sin(\theta_i) \cdot \sin(\phi) \end{aligned} \tag{3.13}$$

The signal model defined in eq. 3.12 is analyzed at $(x, y) = (0, 0)$, and degrades to the 2D monogenic signal in case of i1D signals as (Wietzke et al., 2009):

$$\begin{aligned} f_p &= a \cdot \cos(\phi) \\ f_x &= a \cdot \sin(\phi)\cos(\theta) \\ f_y &= a \cdot \sin(\phi)\sin(\theta) \end{aligned}$$

The 2D analytic signal is built around the local signal representation by tensors, as was suggested in (Johne, 1993), though using as differentiators the generalized Hilbert transforms:

$$\begin{aligned} \mathcal{H}^n f(z) &= \mathcal{H}\left\{\mathcal{H}\left\{\mathcal{H}\ldots\right\}\right\}(z) \\ &= -\frac{n}{2\pi}\int_{\xi\in\mathbb{C}} \frac{f(\xi)}{(z-\xi)^n} ||z-\xi||^{2-n} d\xi \end{aligned} \tag{3.14}$$

where $\mathcal{H}^n$ is the n-th order Hilbert transform, and $z \in \mathbb{C}$.

To explain the notation, we start by defining an analogue to the Hessian matrix (Wietzke et al., 2009) using the second order Hilbert transform:

$$\begin{pmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy} \end{pmatrix}$$

where f_{xx} is the second order derivative in the x direction calculated with the Hilbert

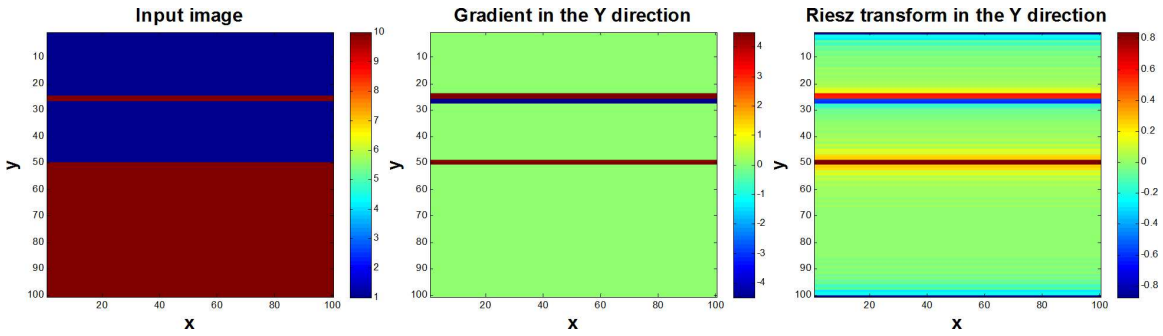


Figure 3.15: Compare the differences and similarities between the Riesz and the gradient response to ideal i1D features in 2-D.

transform as differentiators. Visually the differences and similarities between the Riesz and the gradient response of ideal i1D features in 2-D is illustrated in Figure 3.15. There is a scaling difference but the main features remain the same: we have a peak at the odd edge and a zero-crossing for the even line.

The isomorphism between well-known signal representations and this new signal model is defined as:

$$\begin{aligned} \left([f_b, f_x, f_y]^T, T^e\right) &\cong [f_p, f_x, f_y, f_s, f_+, f_{+-}]^T \\ f_s &= \frac{1}{2}[f_{xx} + f_{yy}] = \frac{1}{2} f_b \\ f_+ &= f_{xy} \\ f_{+-} &= \frac{1}{2}\left[f_{xx} - f_{yy}\right] \end{aligned} \tag{3.15}$$

where the (f_s, f_+, f_{+-}) triplet corresponds to the local signal model defined by the structure tensor in (Johne, 1993) and T^e is the Hessian matrix as defined above representing the even signal component.

In (Johne, 1993)[pg.151] the orientation vector with tensors is defined as:

$$\mathcal{O} = \begin{pmatrix} f_{xx} - f_{yy} \\ 2\ f_{xy} \end{pmatrix}$$

where the subscripts denote partial differentiation. With this notation the local structure is characterised by:

$$S = \begin{pmatrix} \text{phase}(\mathcal{O}) \\ \mathcal{C} \\ f_{xx} + f_{yy} \end{pmatrix}$$

The second component of this structure descriptor is the coherence measure defined as:

$$\mathcal{C} = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \tag{3.16}$$

where λ_i denotes the eigenvalues. This gives the coherence of the local image structure in the direction of the orientation vector defined by $\mathcal{O}$. It returns a value between 0 and 1, where 0 denotes a completely isotropic structure, and 1 a complete anisotropic i1D. The third component of this structure vector gives the certainty of the measurement. This is analogous both in theory and in computation to the dominant component of the local energy given by the response to the bandpass filter.

In the 2D analytic signal notation defined in eq. 3.15 the f_s is the mean square gradient magnitude, f_+ and f_{+-} defines the local orientation vector according to (Johne, 1993) presented above by the $\mathcal{O}$ vector.

The new signal model includes two additional components as compared to the monogenic signal. These are the apex angle and the homogeneous component defined as (Wietzke et al., 2009):

$$\begin{aligned} \alpha &= \arccos\frac{\sqrt{f_+^2 + f_{+-}^2}}{||f_s||} \\ f_h &= \sqrt{\frac{1+\cos\alpha}{2}} \in [0,1] \end{aligned}$$

The apex angle is derived based on the geometry presented in Figure 3.16 and parameterised accordingly. The cosine of the apex angle is based on the triangle defined by the blue, yellow and green arrow that contains α using pythagoras law. The homoge-

neous component is associated with the 3D projective space (Wietzke and Sommer, 2008).

The homogeneous component enables a formal definition of the transition between the ideal i1D and i2D components, while the apex angle quantifies the opening angle between two i1D lines.

From these two features we can derive, for example, that in the case of i1D signals it collapses into the 2-D monogenic signal representation, Figure 3.14. If $\alpha = 0$ then we have an i1D line and $f_h = 1$. We have maximal i2D where $\alpha = \frac{\pi}{2}$ and $f_h = \sqrt{\frac{1}{2}}$.

The features that are also available in the framework of the 2D analytic signal include the mean orientation, the local phase and the local energy, defined as:

$$\begin{aligned} \theta_m &= \arctan \frac{f_h^{-1} f_y}{f_h^{-1} f_x} \\ \phi &= \operatorname{atan2}\left(\sqrt{[f_h^{-1} f_x]^2 + [f_h^{-1} f_y]^2}, f_p\right) \\ a &= \frac{1}{2}\sqrt{f_p^2 + [f_h^{-1} f_x]^2 + [f_h^{-1} f_y]^2} \end{aligned} \tag{3.17}$$

Invariance properties of the 2D analytic signal as compared to the monogenic signal were tested on synthetic images. These tests showed that the 2D analytic signal has a better performance as compared to the monogenic feature estimates despite its shortcoming of not producing an exact zero apex angle for 1D signal. This latter characteristic creates a structure dependent offset, however makes the estimated orientations and associated phase values unreliable, therefore, it is not used in this thesis.

3.2.3 Conformal Monogenic Signal

The conformal monogenic signal (Fleischmann et al., 2011) is one of the most sophisticated current signal models formulated in geometric algebra. This model improves

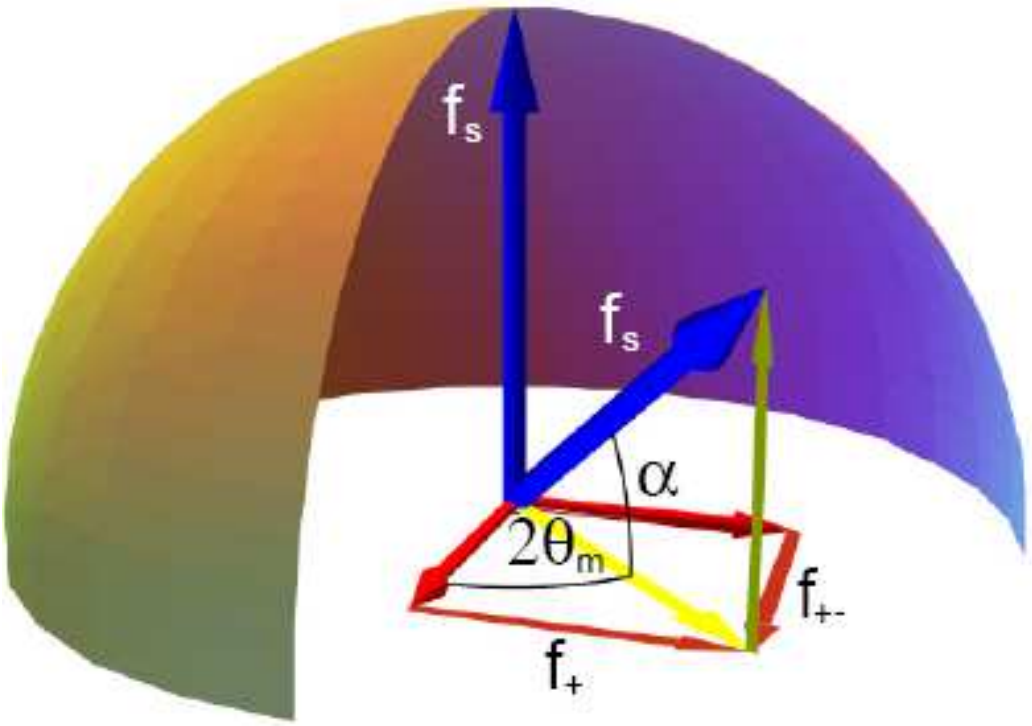


Figure 3.16: 2D analytic signal geometry as was presented in (Wietzke et al., 2009).

upon the 2D analytic signal with a new signal model that includes curved signals as:

$$f(x) = a\,\cos(k|x-m| + \phi) \tag{3.18}$$

where a is amplitude, k is frequency, $m = r_m(\cos(\theta_m), \sin(\theta_m)) \in \mathbb{R}^2$, $x \in \mathbb{R}^2$ and $\phi \in [0, \pi]$. This means that it contains signals that are within a constant radius away from the centre m, which are called isophotes. This can be more formally defined as:

$$\gamma_m(x) = \{y \in \mathbb{R}^2 : |x-m| = |y-m|\} \tag{3.19}$$

with the notation above.

The main design principle behind the conformal monogenic signal is to use elementary geometry that then allows the representation of both i1D and i2D signals within the same framework. Within this framework the 2-D image is sampled by the

local conformal embedding into an $\mathbb{S}^2$ sphere. Both i1D lines and i2D curves are represented with a circle. Since the lines ideally extends into infinity it's representation crosses the North and South pole of the sphere.

Compared to the earlier 2D analytic signal it contains a natural transition between the i1D and i2D signals in form of the isophote curvature, defined as:

$$\kappa_m(x) = \frac{1}{r_m} \tag{3.20}$$

with the above definitions.

The local sampling of the 2-D image is accomplished with a bandpass filter. These samples then are mapped onto a $\mathbb{S}^2$ sphere. In (Fleischmann et al., 2011) the authors propose to position a sphere of radius $\frac{1}{2}$ at the $(0, 0, \frac{1}{2})$ origin, defined as:

$$\mathbb{S}^2 := \left\{ \mathbf{u} \in \mathbb{R}^3 : u_1^2 + u_2^2 + \left(u_3 - \frac{1}{2} \right)^2 = \frac{1}{4} \right\} \tag{3.21}$$

This model allows the definition of the inverse stereographic embedding as:

$$\mathcal{S}^{-1}(\mathbf{x}) = \frac{1}{1 + x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ x_2 \\ x_1^2 + x_2^2 \end{pmatrix}$$

The parameterization of this model is done using the Hilbert transform and follows similar steps to that of the previous analytic signal models.

Given that we do not use it in the current thesis[1] we omit its detailed presentation here. The reason these details are included here at all is because it holds the key descriptor of isophote curvature and it has a natural extension into 3-D, which is important for medical image analysis. As a complementary part to this we report in

[1]Our current experience with this model is that the derived descriptors are highly sensitive to filter scale and do not necessarily reflect the expected numerical curvature values.

this chapter the integral invariant based curvature measurements.

3.3 Phase congruence feature model

The phase values are not everywhere meaningful as feature locations. This fact was realized by (Morrone et al., 1986) (Morrone and Owens, 1987) (Morrone and Burr, 1988) who found that phase values align at feature points over multiple scales, irrespective of whether the feature is an edge or ridge, or whether it is negative or positive. Indeed this observation may be interpreted as a *definition* of a feature, a concept that was previously vague and remains so in much of the image analysis literature. It is reassuring that even the most recent findings in vision research point to the fact that the primary visual system does not as per se encode phase on its one, but instead encodes the coherence of phase over multiple wavelengths (Perna et al., 2008).

As was described earlier there are four values that are meaningful for 1D signal feature detection: positive and negative even features are $k\pi$, odd features are $(2k+1)\frac{\pi}{2}$, where $k \in \mathbf{Z}$ (Figure 3.3).

Phase congruence (PC) is illustrated in Figure 3.17 for an odd step edge and an even triangular pulse. The FT series harmonics align at the centre at the pulse for the even component and at the fall of the edge of the odd component.

Note that all components not only align, but are also in phase at the location of the edge and ridge, and only at that location. This is more obvious from the Fourier series of the step edge which is:

$$FS_{step}(x) = \frac{4A}{\pi} \sum_{k=0}^{\infty} \frac{1}{k} \sin\left(\frac{2\ \pi\ k\ x}{T}\right) + \text{const} \tag{3.22}$$

where $k = 2n+1$, $n \in \mathbb{N}$, A is amplitude, T is period. This meets periodically at each $x = T$.

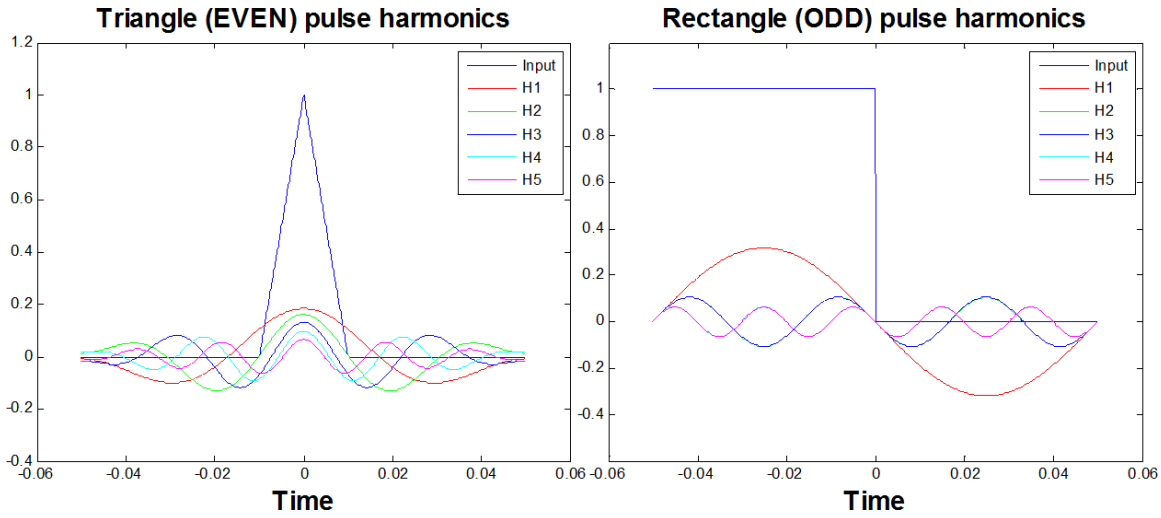


Figure 3.17: Fourier series decomposition of an even triangular pulse and an odd half rectagular step edge. The FT harmonics align at the centre of the pulse for the even triangle and at the fall of the step edge for the odd one.

The idea of PC as a feature measure was not taken up until the energy model was proposed by the same authors (Morrone and Burr, 1988). This means that whenever the local energy of the signal is maximal, we have maximal PC.

There is rationale in this model given that the local energy simply serves as a confidence on the feature itself, however, it is contrast and illumination dependent. For this reason, it is prone to noise. The contrast dependence of the energy as compared to the local phase is illustrated in Figure 3.18.

3.3.1 Kovesi type phase congruency

The next suggestion was to apply the phase congruence weighted by the energy and at the same time to measure the deviation from a mean phase angle at the particular location. Kovesi proposed this method and extended it, by steering, into 2D to be applicable to images (Kovesi, 1999):

$$PC(f) = \frac{\sum_n W(\mathbf{x}) \lfloor A_n(\mathbf{x}) \Delta\Phi(\mathbf{x}) - T \rfloor}{\sum_n A_n(\mathbf{x}) + \epsilon} \tag{3.23}$$

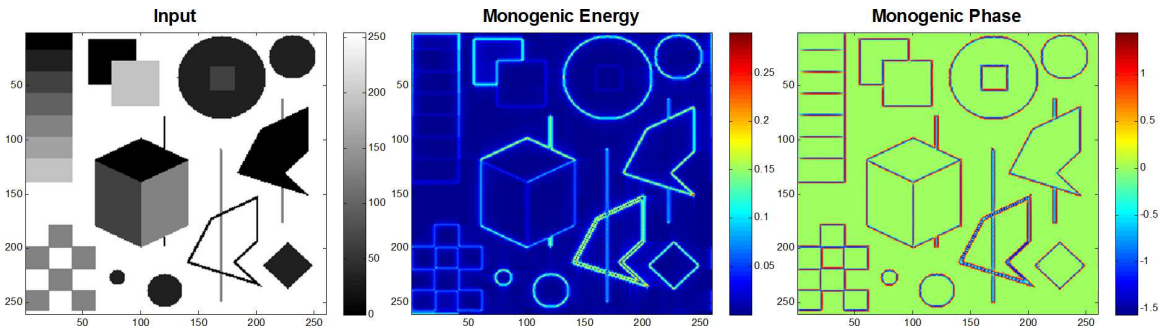


Figure 3.18: Comparison between the monogenic local energy and local phase. Derived using a FT domain implementation with the Mellor-Brady filter, with filter parameters $\alpha = 2.25, \beta = 0.25$ on scale 18. Note the contrast variance of the energy, but not that of the phase information.

with

$$\Delta\Phi(\mathbf{x}) = \cos(\phi_n(\mathbf{x}) - \overline{\phi}(\mathbf{x})) - |\sin(\phi_n(\mathbf{x}) - \overline{\phi}(\mathbf{x}))| \tag{3.24}$$

where ϕ denotes local phase, $\mathbf{x}$ is spatial coordinates, T is a hard noise threshold and A_n is signal amplitude. The weighting function W is a sigmoid function based on the spread of the filter response defined as:

$$\begin{aligned} W(\mathbf{x}) &= \frac{1}{1 + e^{\gamma(c - s(x))}} \\ s(\mathbf{x}) &= \frac{1}{N}\left(\frac{\sum_n A_n(\mathbf{x})}{\epsilon + A_{max}(\mathbf{x})}\right) \end{aligned} \tag{3.25}$$

In 2-D eq.3.23 is extended and summed over the different orientations in addition to the number of scales n. This is not an isotropic extension of the phase congruence, but it was a reasonable solution given the lack of an isotropically extended Hilbert transform to compute the real 2-D isotropic conjugate signal pair in form of the Riesz transform. Additionally, we also lacked the specific signal model that would have allowed the derivation of the local image features as was proposed in (Felsberg and Sommer, 2002).

3.3.2 Reisfeld and Schenk type phase congruency

A constrained phase congruency transform (CPCT) was also proposed in (Reisfeld, 1996), which simultaneously detects interest points along their scale and orientation. To do this it uses a binary coding of four ideal feature types, namely the +ve/-ve odd and even features.

Based on the previous PC considerations and the observation that it is simply not correct to calculate the PC over an arbitrary number of scales a new algorithm was proposed to calculate the PC. (Schenk and Sir Brady, 2003) selected prospective feature points based on a pre-filtered set of high local energy regions. The stability of phase is traced based on consideration to avoid the inclusion of scales where the local phase is corrupted by the presence of nearby structures and/or is gathered by overlapping filter responses.

3.3.3 Complex phase order type phase congruency

A more recent complex phase order (CPO), or phase congruency, measurement was put forward in (Wong et al., 2010). This model uses a cascaded noise filtering approach. First, it proposes to apply a local signal adapted kernel during the derivation of the quadrature filter responses, and second, it uses the Kovesi type PC model (Kovesi, 1999) with filtered energy weights. It uses a bilateral filtering approach to differentially describe the noise and the real signal statistics of the underlying measurements. However, by experimenting with the proposed PC it turns out that this method often associates a strong weight with the homogeneous tissue components.

The CPO is derived based on the quadrature pair:

$$\begin{aligned} f_{even} &= f(\mathbf{x}) \star K \star R_{s,\theta}^{even} \\ f_{odd} &= f(\mathbf{x}) \star K \star R_{s,\theta}^{odd} \end{aligned} \tag{3.26}$$

where $\star$ denotes spatial convolution, $R_{s,\theta}$ is the quadrature bandpass filter (they used Log-Gabor) and the adaptive kernel is defined by:

$$K(\$) = \exp\left(-\frac{\sigma(\mathbf{x})^2(f(\mathbf{x}) - f(\$))^2}{\zeta\ \sigma_b^4}\right) \tag{3.27}$$

with $\$$ being a sampled set around $\mathbf{x}$, σ_b is the standard deviation(STD) of the background and σ is the STD from a small neighbourhood around $\mathbf{x}$, ζ is given as a constant. In practice there is no detail provided on how each of these are estimated. Based on the filter responses the eq. 3.23 is then used to derive the PC with the weight term being re-defined as:

$$W(\mathbf{x}) = W_\nu(\mathbf{x})W_\tau(\mathbf{x}) \tag{3.28}$$

where the different weights are constructed based on the expected signal a noise behaviour over scale as was proposed in (Field and Brady, 1997).

All of the above approaches rely on the local energy of the image as a measure of local phase certainty. However, the energy is contrast dependent therefore it does influence the accuracy of the proposed algorithm. A comparison between three PC approaches (Riesz-weighted PC, Wong PC and Kovesi PC) on breast MRI shows their performance, Figure 3.19.

With the note that within our experiments we avoided the application of the adaptive kernel in the Wong type PC implementation we draw the following conclusions. (All methods are implemented using the Log-Gabor bandpass filter.) The Kovesi type PC does only detect the breast boundary, but does not supply information on the inner texture of the breast. The bilateral filtering based PC does perform better than the original Kovesi type PC, but it does give a high weight to tissue regions that are relatively homogeneous. These are the regions, for example, close to the surface of the breast. As opposed to these, the Riesz-weighted PC, that is proposed in this

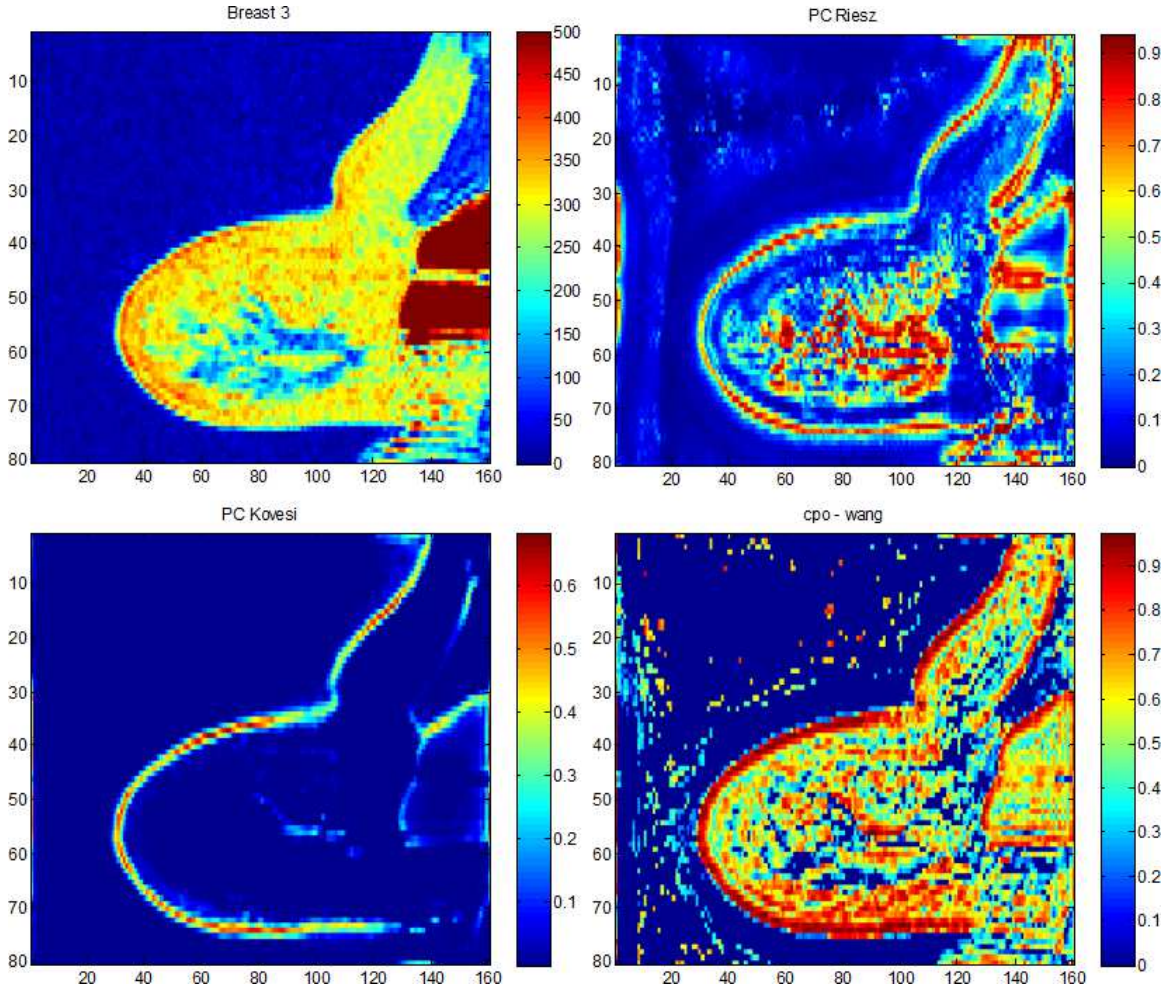


Figure 3.19: Comparison between three phase congruency estimates applied to breast MRI. Clockwise: input breast image, Riesz-weighted PC, Wong bilateral-filtering based PC and the Kovesi type PC.

thesis, does highlight the duct system structure of the inner breast and it does not give any weight to the homogeneous tissues.

In what follows we present a way to derive isophote curvature based on integral invariants and to simultaneously select scale and feature within an isotropic non-parametric scale-space.

3.4 Integral invariant for curvature approximation

Geometric invariance has been recognized as an important tool in computer vision and medical image analysis. Its main application is to extract features for recognition, matching, reconstruction and visualization (Hong, 2004). Integral invariants are robust to noise, and embody an inherent scale parameter through the radius of the defining circle, from the smallest scale r_0 to the maximum scale r_{max}.

As was pointed out earlier, curvature is an example of invariance under the group transformations of rotation and translation. It provides a unique local characterisation of a curve. However, its computation typically relies on the computation of second order derivatives.

The uniqueness of curvature brings the completeness of the representation with the possibility to recover the original signal up to a symmetry transformation. As a result, all the differential invariants studied in image analysis are a function of curvature. The measurements are based on the general diffusion scale-space through filtering with a Gaussian kernel. This means that the extremae that are of interest for feature detection become displaced. By retaining the advantageous properties of differential invariants and improving on those drawbacks we arrive at the definition of the integral invariants, with specific focus on the Area Integral Invariant.

Define the characteristic kernel $\chi(B_r(p), C)(x)$ as (Hong, 2004):

$$\chi(B_r(p), C)(x) = \begin{cases} 1 & \text{if } x \in B_r(p) \bigcap \overline{C} \\ 0 & \text{otherwise} \end{cases} \tag{3.29}$$

where, $B_r(p)$ is the measurement disk that weights each $p \in C$ point and $\overline{C}$ is an interior region in curve C. Then the local integral invariant $I_r(p)$ at each point p is

defined as

$$I_r(p) = \int_{\Omega} \chi\left(B_r(p), C\right)(x)dx \tag{3.30}$$

Aside from robustness to noise, integral invariants have an inherent associated scale-space in the guise of the radius of the measurement circle. Similar to differential invariants, the completeness of the representation is satisfied; i.e. a given shape can be reconstructed from its integral invariant signature. Although the defined decomposition does not in the strict sense form a scale-space, its extrema are not dislocated, assuming that just one feature exists in the measurement window.

The inherent scale component in form of the radius of sampling window makes it possible for integral invariants to be normalized to the area of the kernel, which may provide a scale-invariant description of the curve.

Finally, the derived area integral measure may be related to the differential invariant measures through the curvature as follows:

$$I_r(p) = r^2 \text{acos}\left(\frac{1}{2} r\kappa(p)\right) \tag{3.31}$$

Where p, r, κ denotes the spatial position, radius of the spacing window and curvature respectively.

3.5 Joint feature-scale saliency

Several algorithms have been developed to detect features on multiple scales. In general, the selection of significant features using scale-space approaches is based on the selected scale, maximizing SNR.

The 'trouble in paradise'(phrase quoted from (Selesnick et al., 2005)) derives from the assumptions. Scale and salient feature selection poses a circular question. That

is, the scale cannot be selected without knowledge of the salient features; conversely, the salient features cannot be defined specifying the scale. This challenge is addressed by the scale-saliency algorithm (Kadir and Brady, 2001), which is a generic model of feature saliency that measures saliency both over scale and feature-space.

Saliency means rarity, or local complexity/unpredictability of the region. This is also responsible for the pop-out effect according to Julesz's definition of pre-attentive vision; this local measurement may only take relative values, because it is inherently defined by the corresponding measure.

One established method is to measure local complexity via Shannon entropy based on an approximation to the probability distribution in a local region (Gilles, 1998). Analogous to the main concept put forward in this subsection regarding integral and differential based methods, the detection of interest regions as opposed to interest points is suggested. One shortfall of Gilles's method, however, is that it uses pre-selected global and local scales with no particular attention paid to the inter-scale changes in the local entropy measures; this is overcome by the suggested scale-saliency algorithm (Kadir and Brady, 2001).

At the heart of scale-saliency lies a key observation and a key idea. Entropy measures disorder based on the local pdf, which however discards spatial information. This means that features with different local structures may lead to the same entropy descriptor. However, the authors suggest that the change of entropy between scales is unique to the fundamental local shape, for example, lines and blobs. The peak in the inter-scale change in entropy measurement defines the most salient scale, see Figure 3.20.

The authors' key contribution is to use both the inter-scale saliency measurement and the features selected based on local entropy measurement on a single scale to

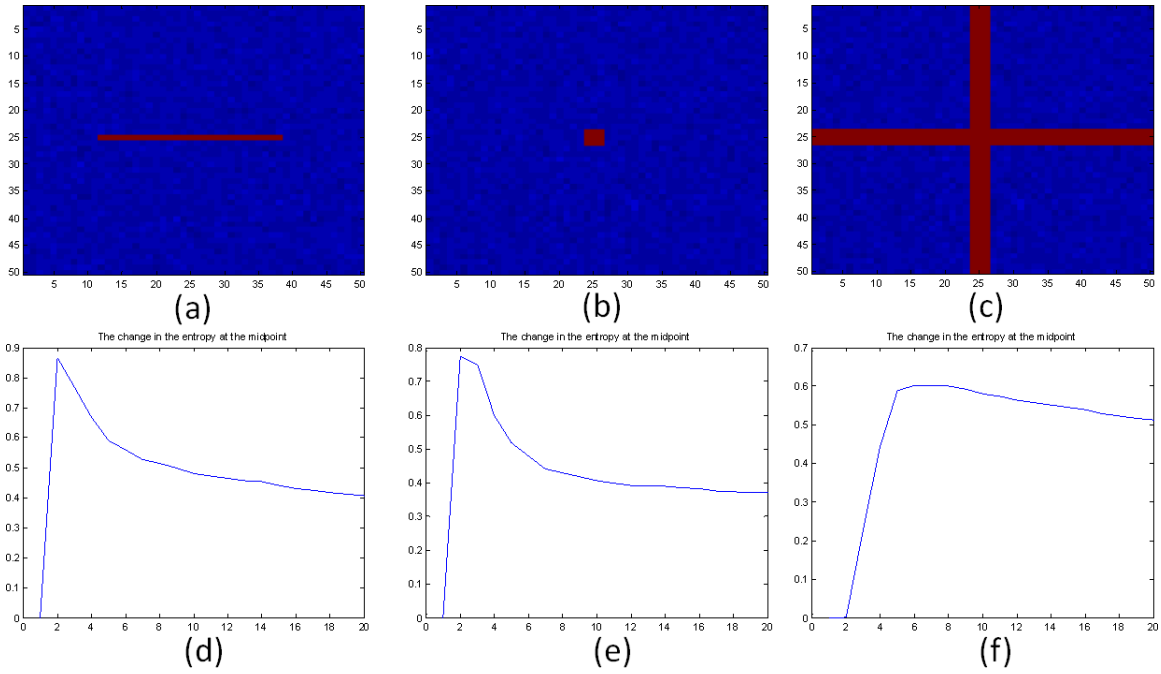


Figure 3.20: Structural features and associated local entropy changes over scale.

define the scale-saliency map as the product of local entropy and inter-scale saliency:

$$\mathcal{Y}_D(s_p, x) = \mathcal{H}_D(s_p, x) \times \mathcal{W}_D(s_p, x) \tag{3.32}$$

with

$$\begin{aligned} \mathcal{H}_D(s, x) &= -\int_{d \in D} p(d, s, x) \log_2 p(d, s, x) \mathrm{d}d & (3.33) \\ \mathcal{W}_D(s, x) &= s \int_{d \in D} \left| \frac{\partial}{\partial s} p(d, s, x) \right| \mathrm{d}d & (3.34) \\ s_p &= \left\{ s : \frac{\partial^2 \mathcal{H}_D(s, x)}{\partial s^2} < 0 \right\} \end{aligned}$$

where, $\mathcal{Y}_D$ is the proposed saliency measure, $\mathcal{H}_D$ is the Shannon entropy, and $\mathcal{W}_D$ is the inter-scale saliency measure, as a function of position x, scale s and the feature d that takes values from D. An example computation may be seen in Figure 3.21.

As opposed to the sampling method presented in (Smith and Brady, 1997), the scale-saliency framework offers the possibility to build a model-free scale space which

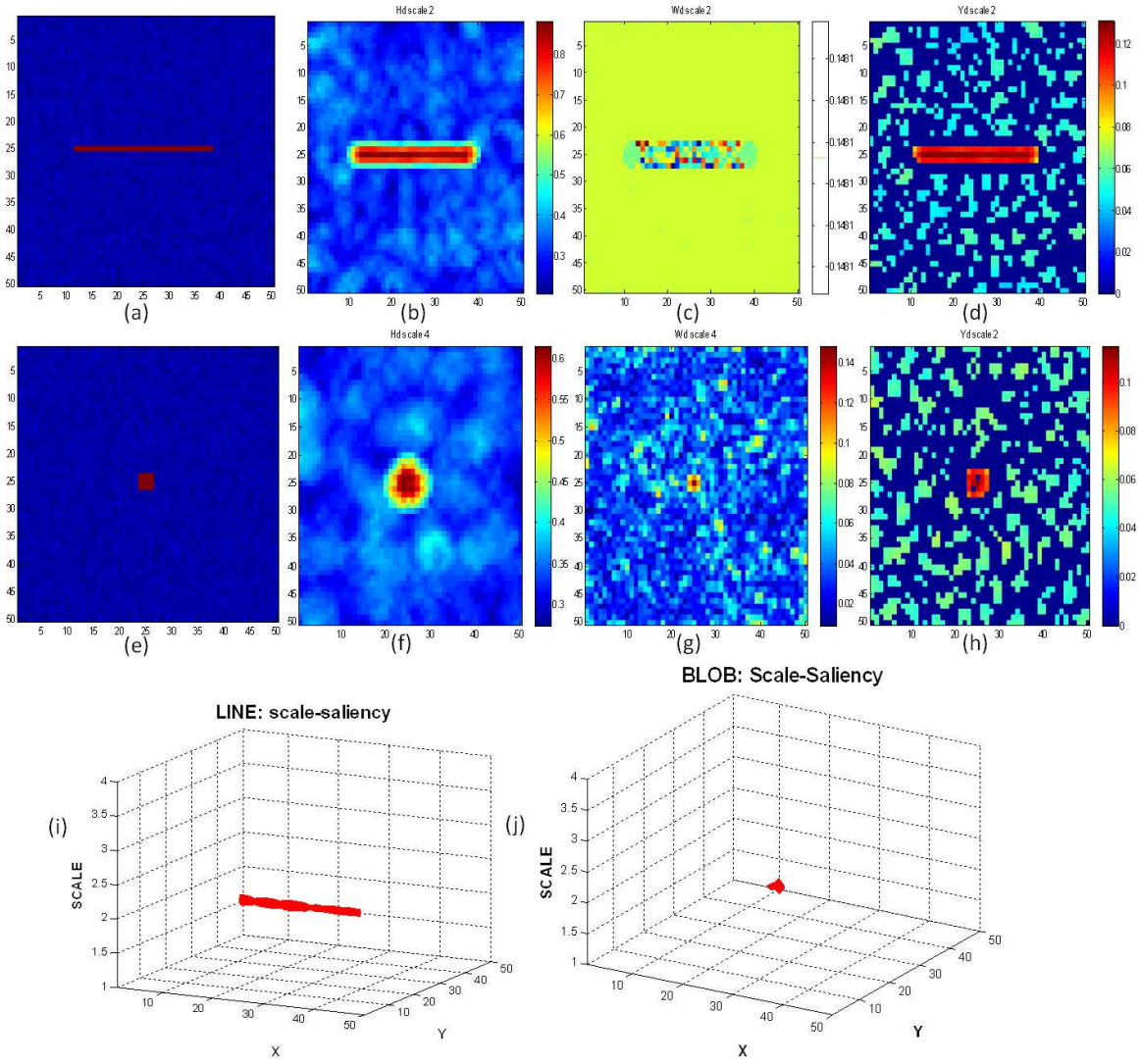


Figure 3.21: Examples of scale saliency maps are shown for an ideal line and blob. The images show the input images (a-e), the local entropy maps (b-f), inter-scale saliency maps (c-g), and scale-saliency maps (d-h) on the selected scales. (i-j) illustrates the respective scale saliency maps after a grouping and thresholding procedure was applied as suggested in (Kadir, 2002)

is either isotropic or fully affine. As was stated earlier, the isotropic scale space is fully determined by the radius of the sampling top-hat kernel; the sampling kernel of the affine model takes the shape of an ellipse and is parameterized by three scale factors: the major axis, minor axis and the rotation angle. This construct fully describes the local structure as a function of structure strength and orientation. The drawback of the latter approach derives in practice from the considerable increase in

the computational time.

Apart from salient feature detection, the method was shown to perform well in characterizing structural textures. The texture description is based on the assumption that the scale-saliency values have good descriptive power for the selected salient scales. Several approaches may exist for texture characterisation; particularly good discriminative power was achieved with the joint saliency and scale histograms for the Brodatz texture database (Kadir, 2002).

The derived scale-saliency features satisfy most of the requirements set by computer vision and image processing tasks. These include robustness to geometric and photometric transformations, in-plane rotation, spatial scaling, translation and intensity scaling. Additionally, it achieves good repeatability and intra class variance under local perturbations.

The repeatability of the method for the isotropic scale-saliency algorithm was improved in several aspects in (Shao, 2005). The accuracy of the scale-saliency maps are based both on the stability of the local entropy and the inter-scale saliency maps $\mathcal{W}_D$. $\mathcal{W}_D$ is based on a differentiation operation, therefore its stability is affected by noise. Several approaches to increase this stability were tested, ranging from bin interpolation to a more accurate approximation to the inter-scale map. The biggest impact on the repeatability measure, resulting in a 25% increase, was achieved by a sampling strategy of the local area based on the observation that the central part is more reliable and the difference between scales increases with scale, i.e. W_D becomes less reliable at lower scales. As a consequence, the weighted count of the pixels within the central region of the sampling window was suggested; for example for two consecutive scales s and $s-1$ count the pixels within a radius $s-1$ twice and count the pixels that fall between s and $s-1$ only once.

In practice there are several challenges with the anisotropic method, one of which is its time complexity. The computationally most expensive solution is the exhaustive

search within the parameter space to find the best matching shape of the kernel to the local structure. One further consideration is the local pdf measurement. The sensitivity of local histogram based measurements is outperformed by Parzen window and NP windows (Joshi, 2007) based estimation methodologies. These, however, add substantially to the computation time. A third time-consuming action is the computation of the local masks; whether this could feasibly be pre-computed is questionable, given the complexities that may arise from an arbitrary local structure.

The method inherits all the advantages of integral based scale space approaches: robustness to noise and an inherent model free scale space.

3.6 Strengths and weaknesses of these approaches for the liver images

We have implemented the different analytic signal models and evaluated the different features available for ideal and liver MRI images, and we have reached a number of firm conclusions. For example, not all monogenic features are useful for the liver images. The local phase (LP) and feature asymmetry (FA) provide a great detail of the structures that describe the internal texture structure of the liver and makes the blood vessels and ducts visible. We compared the 2D monogenic signal (Felsberg and Sommer, 2002) to the 2D analytic signal (Wietzke et al., 2009) and the conformal monogenic signal (Fleischmann et al., 2011). We concluded that the local orientation for the latter extended approach are more accurate than the original monogenic signal that only includes i1D features. However, the local phase values are less reliable therefore we chose to use the 2D monogenic signal in all of our experiments with the liver MRI images. (Note however, that the local orientation is an important element of the ciD framework.)

Local phase gives a reasonable structural description of the image features in a

rotation invariant manner. However, the LP contains the second norm of the Riesz components, so it loses information that is inherent in these components. These will be discussed in detail in the next chapter.

3.6.1 Sensitivity to noise

The Hilbert and Riesz transforms are differentiators with all-pass filter characteristics (Figure 3.15). Most of the noise removal property is dealt with by the even bandpass filter. Filters that have small support are less prone to noise effects, or maybe more influenced dependent on the relative size of the window as compared to the feature of interest. While filters covering big areas will accumulate noise, at the same time they potentially perform a local averaging and hence suppress noise. There is no one design principle behind an optimal bandpass filter for 2D analytic signal model.

Alternative (to differential based) methods were presented to set-up a non-parametric isotropic feature scale-space and derive local curvature based on integral invariants. As the name of the latter one suggests this performs an integration instead of differentiation, which is very robust in the presence of noise. At the same time a model-free isotropic scale-space is also desirable for medical image analysis, but we note that the specific simultaneous feature and scale localisation gives an overly sparse representation and clutter removal forms an integral part to identify significant features.

In general, the monogenic signal is less prone to noise effects as compared to tensor based methods, though these are more prone to noise effects than their integral invariant counterparts. This is because the tensor based methods use high-pass differentiators that amplify the noise effects, the monogenic signal uses all-pass differentiators, while the latter approach uses no differentiation at all.

3.6.2 2-D Analytic Signals and i2D structure

The original 2-D monogenic signal (Felsberg and Sommer, 2002) is strictly only applicable to, and the derived features only have fidelity for, i1D signals. To alleviate this shortcoming, the monogenic structure multivector was suggested (Felsberg, 2002).

Later developments that are based on evolved signal models incorporate this shortfall directly. The 2D analytic signal (Wietzke et al., 2009) suggests that the apex angle as an i2D feature to be used with the assumption of two intersecting i1D lines.

As opposed to this, the conformal monogenic signal (Fleischmann et al., 2011) makes the characterisation of the transition between i1D and i2D explicit through the isophote curvature. This is possible because the signal representation itself is composed of curved planar waves. Overall, it is always the fundamental signal model that changes with each approach. We note that the conformal monogenic signal's curvature is extremely sensitive to the scale of the filtering.

3.6.3 Concluding remarks

We have presented a brief review of the developments from the 1D analytic signal into higher dimensions and two alternative feature detectors. Each development of the analytic signal is based on one idea, Figure 3.7, 3.16:

1. Expand the earlier signal model to include the desired signal geometry, while preserving the earlier one;

2. Lift the new signal model into a higher space, e.g. projective, conformal;

3. Use linear operations to derive the desired signal features.

The conformal monogenic signal is the most recent signal model that includes local analysis of both i1D and curved signals with an exact curvature calculation with all the advantages of rotation invariance. The scale-saliency algorithm allows the

detection of structural features in an isotropic or affine scale-space by concomitant maximization of featureness within and across scales. In addition we have shown that the method suggested in (Hong, 2004) provides the derivation of curvatures using integral invariants that are robust to noise.

Chapter 4

The Riesz transform and shortcomings of the monogenic signal

In this chapter we investigate a number of previously unremarked shortcomings of the monogenic signal, even in case of simple input signals. On the other hand, we demonstrate the power of the Riesz components and the Fisher Z transform in extracting information about textures in images.

4.1 Introduction

We propose a novel approach to the quantification of feature energy/certainty by the concept of phase congruency (PC). We replace Kovesi's energy weighted PC by an approach that seeks to put texture at the core of the idea. Subsequently, we illustrate the approach on the application that motivated us in the first place: liver Magnetic Resonance (MR) images in which the goal is to assist a clinician to characterize different stages of liver disease.

Our approach has its basis in two ideas: the first concerns the limitations of the

Monogenic Signal (MS) as it currently stands; and second is the realisation that the Riesz components are powerful elements for texture description. All the in-silico experiments reported here have been designed to illustrate these two main ideas. In the following paragraphs, we briefly summarize the hypotheses and conclusions of those experiments.

First, we conduct a series of experiments to determine whether textures may be decomposed into their component parts using the Riesz maps. We conclude that there is a rich source of information in the two individual Riesz maps and that often (almost always) it is worth decomposing the image into finer structures that aid interpretation, recognition, and understanding of the texture. None of the MS components are able to deliver this information (except perhaps the local phase).

There is one limitation to the Riesz components: the features they provide are differentials and they are not rotationally invariant. One result that we have is partly due to this observation and to recognising that the local phase (LP) is rotational invariant and depends on the second norm of the Riesz components. The "problem in paradise" starts because LP does not preserve all the information contained in the individual Riesz maps. This observation leads us to the idea of Riesz map weighted LP.

The Riesz map weighted LP, and then later phase congruency (PC), may be considered to be a version of the Kovesi type phase congruency; but, crucially, we do NOT use the energy to weight the phase values. Instead we use the absolute values of the Riesz components. By experimenting with different forms of the absolute Riesz components we conclude that the Riesz weighted LP gives results that reveal details that are hidden even to the general energy weighted LP congruence. The results of this novel process are then demonstrated both for general computer vision images and for different stages of liver disease.

The combined use of the Riesz maps (fu and fv, which are the Riesz components in

the x and y direction) and the LP was investigated from the standpoint of grouping. We evaluated the use of phase and fu, fv using k-means and mean shift in two ways on multiple scales. When we use three consecutive scales and a particular point-wise grouping technique, the mean-shift produces a grouping that looks like a phase congruence (PC) map of the Smith test image (Kovesi, 1999).

To evaluate the power of local entropy in texture characterisation we evaluated the use of the conventional histogram based and the NP-window computed entropy on computer vision test images, Randen and Husoy texture mixtures, and different stages of liver disease. Our results indicate that using the local pdfs leads to a very promising approach to capturing, and quantifying, texture information.

Following exactly the same process, we evaluated and compared the Riesz weighted PC with the energy weighted one proposed by Kovesi. We found that the energy weighted measure captures more of the boundary but much less of the texture and vice versa with the Riesz weighted PC. We also found that fractal measures computed on these PC images produces results that shows a consistent trend for the Riesz weighted results with the different stages of the liver disease.

The Fisher Z transform with cross-scale correlation and hypothesis testing has also proved to be an invaluable source of structural features.

All of the above results reinforce our contentions about the power of the (fu, fv) Riesz pair on multiple scales. In summary, the Riesz weighted PC detects more of the structural changes that may subsequently be used for the quantification of the stage of the liver disease.

4.2 Riesz components

The two major shortcomings of the monogenic signal (Felsberg and Sommer, 2002) are: first, that it *fundamentally and intrinsically* provides only a single orientation,

namely the ratio of the two Riesz components; and second, that the local energy has no meaning in many practical cases. In more detail, the LO becomes distorted if the values in either of the Riesz maps becomes either too large or too small. However, we have found that information with respect to multiple orientations may instead be obtained using the Radon transform which also provides information about frequency in itself. This is illustrated in Figure 4.1 and 4.2.

The first image shows a scene after a single "pebble" is thrown in a calm pool of water and the generated circular map that results from this. The figure contains the Riesz components in the second column, the monogenic components in the first and second column and the Radon transform in the third column. It may be observed that the energy (in blue, first column) provides little or no information about the original structure. It may also be observed that the Radon transform contains a continuity of orientations - that is due to the circle - at approximately equal distances, that are the frequencies due to the separation of the concentric circles.

Figure 4.2 takes this one step further. It is a simulation of a situation in which two "pebbles" are dropped simultaneously into a calm pool. We are in this case interested in their interaction: can it be resolved? Can the fact that there are two interacting (simple) patterns resolvable? The generating pebbles at different locations are also shown, Figure 4.2. Again, not much information is provided by the LE. However, there are substantial changes detected by both Riesz components. On the other hand, we can discern a great deal of information from LP, LO and the Radon map. The Radon map shows disrupted lines which corresponds to the disrupted orientations of the input image. The different frequencies are also visible. How the exact, quantitative read-out is derived from the Radon map is not obvious, nor is that our primary concern. Rather, it is obvious that the LP represents the original pattern, and it is evident why the orientation map is broken at those parts where it is broken in the input. In conclusion, in this example, all maps apart from the

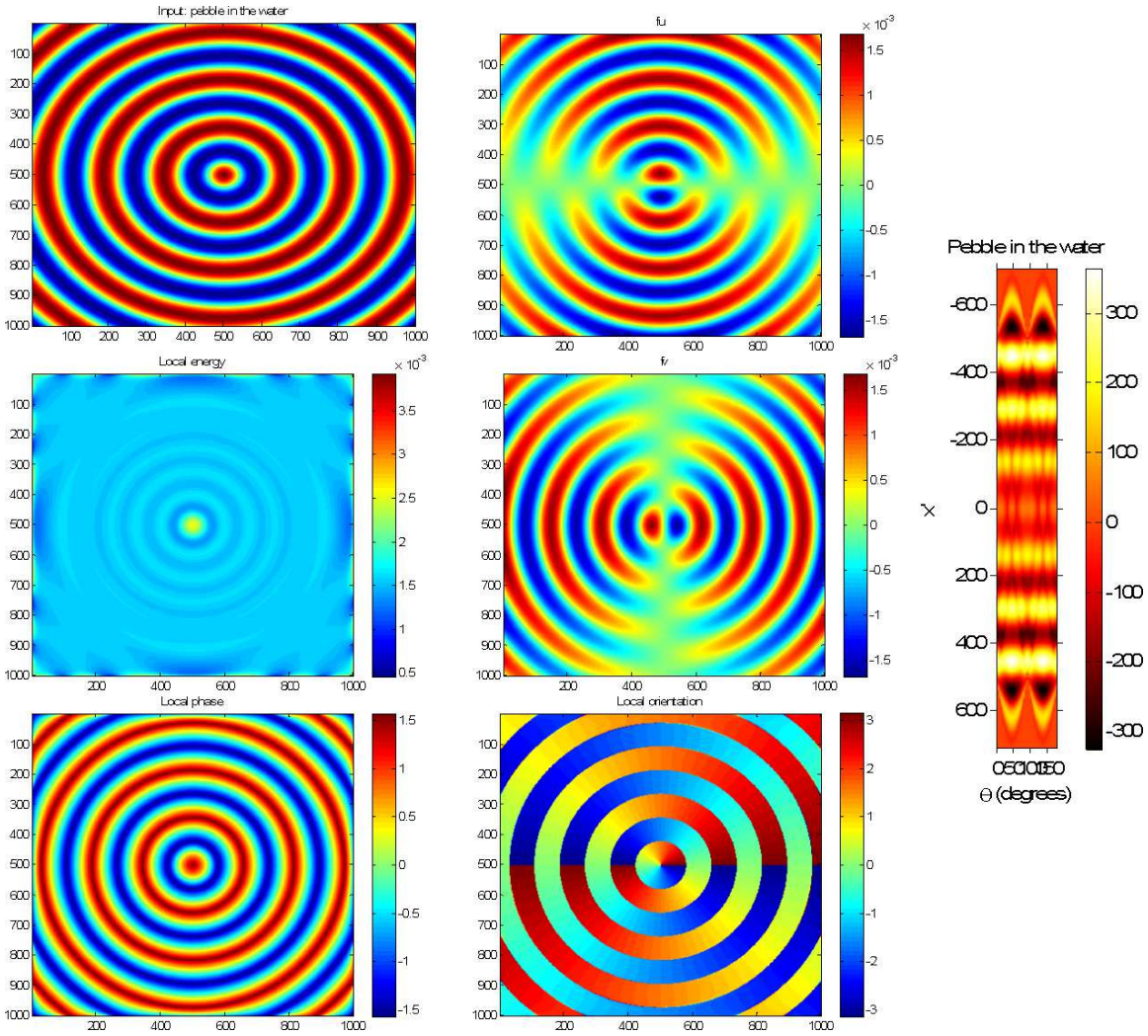


Figure 4.1: Single pebble and its components. The fu, fv denotes the usual Riesz components. The rest of figures as labelled are the local energy, phase and orientation. The right hand side orange figure shows the Radon transform. Note that both the different orientations and frequencies are present. This will be clearer in the next figures.

energy show great promise for further processing. The pair of "pebbles" shown in the above experiment is illustrated in Figure 4.3. Next, the centre of the second pebble is moved in the orthogonal direction. Again, the figure shows the generating components, the monogenic signal, the Radon transform and the (fu, fv) pair. This latter one shows the structures in both orthogonal directions. Again, the Radon map shows the details, and once again the local energy provides no useful information.

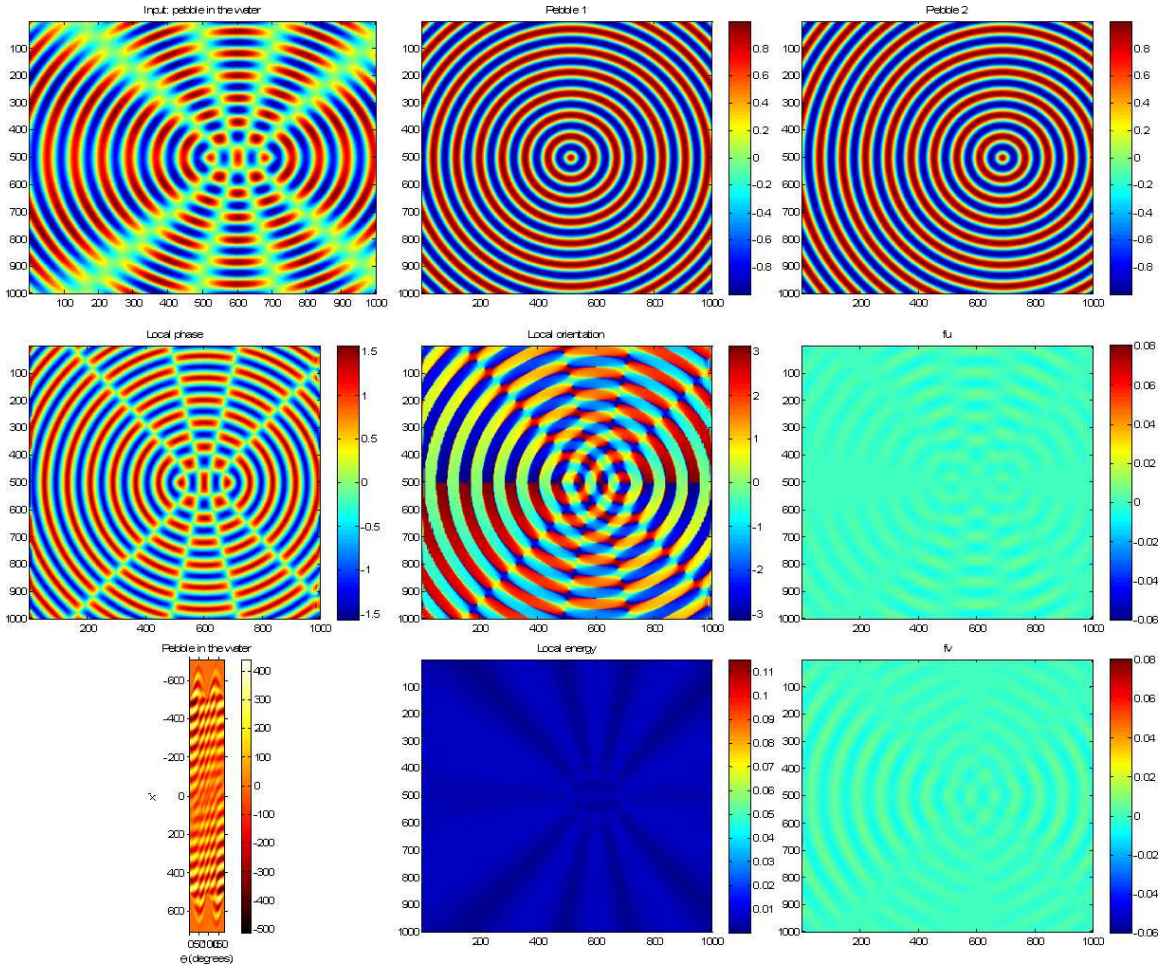


Figure 4.2: Two pebbles and their superimposed components. The fu, fv denotes the usual Riesz components. The rest of figures as labelled are the local energy, phase and orientation. The left bottom side orange figure shows the Radon transform. Note that both the different orientations and frequencies are present.

We conclude that the Riesz components are a source of substantial structural information that are in many cases more valuable than simple gradients because their source is all-pass as opposed to high-pass filters and, as a result, the responses can be finely tuned to the necessary one using an appropriate bandpass filter. We found that using the Riesz maps has several major benefits as opposed to the monogenic components:

1. they provide structural information that is lost during the LP calculation; and

2. they may be used to decompose to an extent (as much as possible) simple

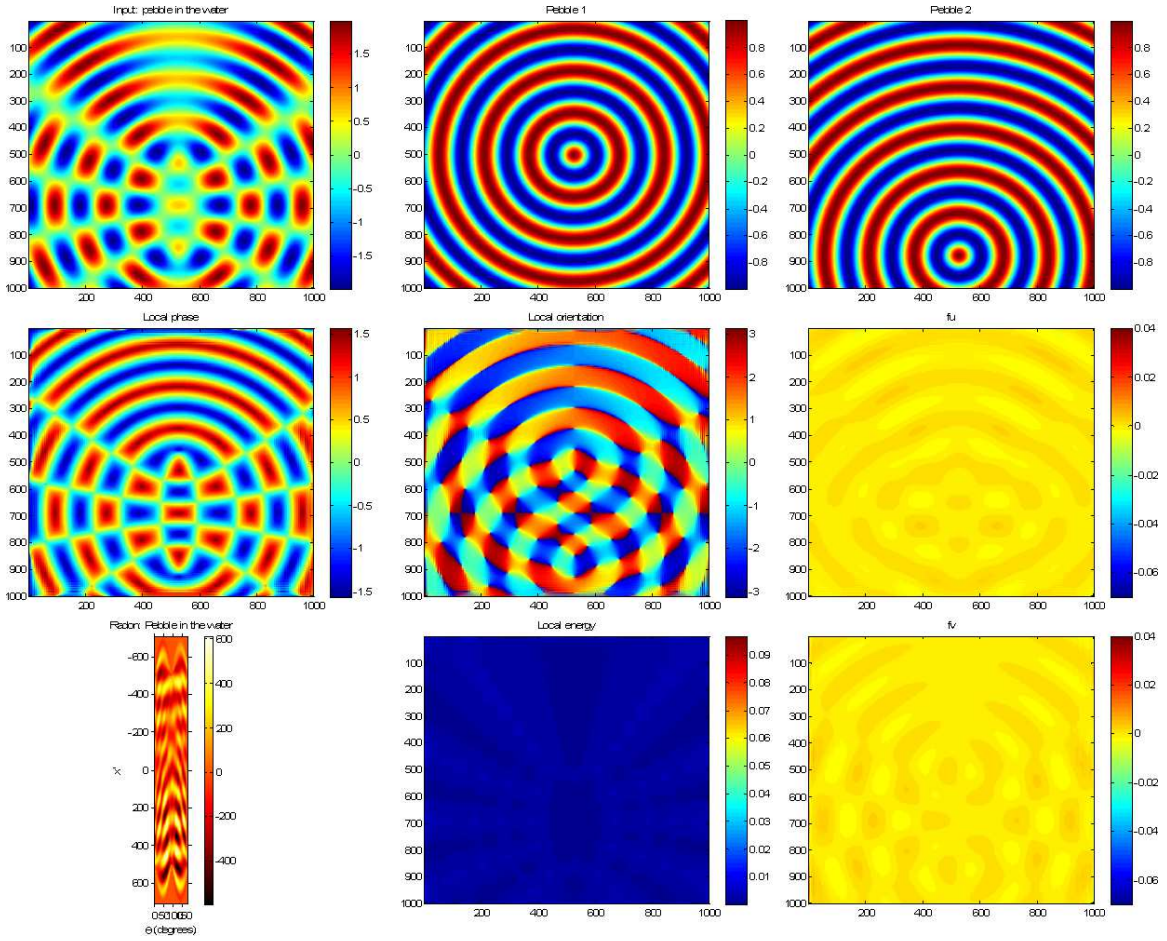


Figure 4.3: Two pebbles and their superimposed components displaced in the y direction. The fu, fv denotes the usual Riesz components. The rest of figures as labelled are the local energy, phase and orientation. The left bottom hand side orange figure shows the Radon transform. Note that both the different orientations and frequencies are present.

oriented textures.

4.2.1 Decomposition of textures

Second, we show that this formulation of the Riesz transform may be used to decompose textures. We provide an example using two superimposed different oriented sinusoid patterns. We use these patterns to illustrate that the (fu, fv) pair is capable of representing more of the intrinsic texture than the phase (though again we note that the latter is rotationally invariant) or any of the individual monogenic compo-

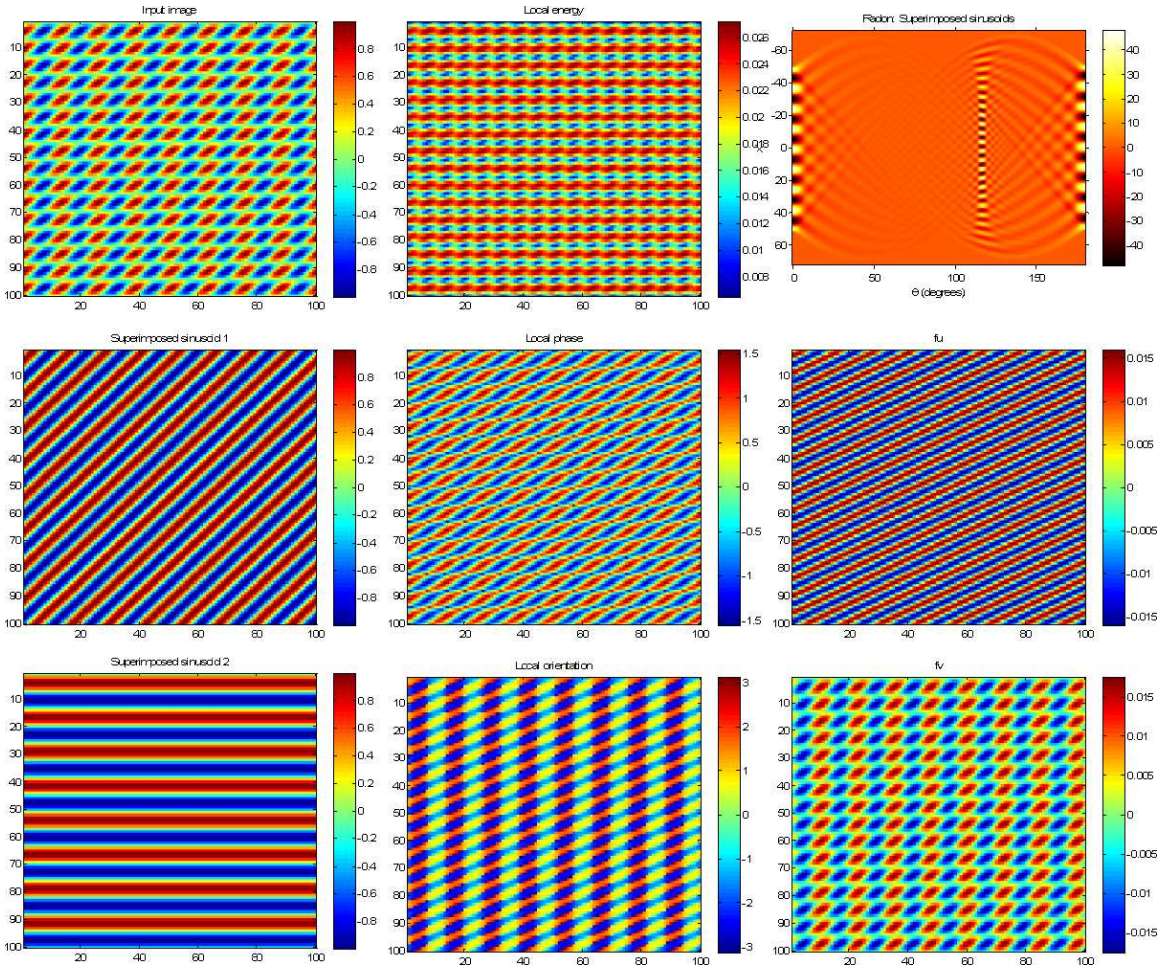


Figure 4.4: Decomposition with the Riesz transform. Two superimposed sinusoids. The fu, fv denotes the usual Riesz components. The rest of figures as labelled are the local energy, phase and orientation and the generating components. The right hand side orange figure shows the Radon transform. Note that both the different orientations and frequencies are present. Also note the decomposition power of the Riesz components.

nents. We provide examples of the Riesz components, as well as of the monogenic and Radon maps. This capability of decomposition of the Riesz components is illustrated in Figure 4.4.

The collection of images shows in the first column the generation of two sinusoid patterns (left centre; left lower) that are then superimposed to produce the input image (the left upper image). The second column presents the monogenic components. Finally, the right column shows the Radon map and the (fu, fv) pair. We immediately

observe two things: first, that the Radon map represents both the two different orientations and the frequency of the generating figures (given in the last column); and second that the (fu, fv) pair gives information about the original sinusoids. At least fu has an orientation that is similar to one of the patterns. On the other hand, neither of the monogenic components (local energy or local orientation) makes this information explicit.

We conclude that the (fu, fv) pair appears to show promise for decomposition of the textures into their intrinsic patterns. Although the LP contains the second norm of the Riesz maps it does not preserve the properties of these components. This is why we suggest that any texture quantifier must include, in a way yet to be made explicit, the Riesz components.

4.2.2 Grouping using multi-scale Riesz components

In this context, grouping means to assemble feature locations, similar to what we do during image segmentation. In general, grouping reduces to selecting certain features to compare, and a comparison measure. Different segmentation methods perform this differently. If we were to analyse coloured images, then we would want the same coloured items to be in the same group. One version of the so called level-set algorithm (Caselles et al., 1997), for example, uses two kinds of information to segment the area of interest.

One of them is the edge information that stops the level set from overspill; the second is the pdf of the area that describes the distribution of the pixel intensity levels. This is just one form of the segmentation algorithm. There exist several features that may be used. For example, phase values, orientation values, texture measures, edge/ridge measures.

One particular example of edge/ridgeness measure is the Inter Level Product (ILP) defined in (Anderson, 2007). It uses phase information on two consecutive scales to

define feature points. It uses knowledge of the octave filters (in this case the Cauchy) (Anderson, 2007) to define how we can nullify the effect of the filter's offset which is generated by the centre frequency. Most important however is that with octave filters each finer scale rotates at twice the rate of the coarser scale. This determines how fast one should rotate away from a feature point. The phase is expected to change linearly with scale around a feature. We also require that the feature be isolated. We note that other features in the vicinity will inevitably disturb the phase values.

The ILP leads to a definition of the phase of a feature similar to that of phase congruence. With the above defined properties we expect the finer scale response at δd away from the feature to be:

$$F_s(\overline{x}) = A \cdot e^{i(\theta_f - c*\delta d)} \tag{4.1}$$

where F_s is the response on the scale s, A is amplitude, $\overline{x}$ is the spatial position, θ_f is the feature phase and δd is the displacement, while c is the particular offset dependent on the filter response measured in rad/spatial unit. With this definition the finer scale response becomes:

$$F_{s-1}(\overline{x}) = B \cdot e^{i(\theta_f - 2c*\delta d)} \tag{4.2}$$

Taking the double of the coarser scale and multiplying by the conjugate of the finer scale we get the feature phase. That is:

$$\begin{aligned} F_{s-1}^{ILPfeature} &= \overline{F_{s-1}} F_s^{double} \\ F_{s-1}^{ILPfeature} &= ABe^{i\theta_f} \end{aligned} \tag{4.3}$$

In what follows we do not use the ILP. Rather, we leave the world of filtering in order to consider methods for grouping based on computational properties of the

images. Although this is not as sophisticated as the grouping used in the statistical data processing it satisfies our purposes and demonstrates the point we wish to make.

The discriminative power of the (fu, fv) pair and the LP on multiple scales was tested using multiple approaches. Grouping means that the inter cluster distance is maximized and the intra-cluster distance is minimized. That is, we aim to find those groups that are closest to a group centre while placing the second, or other groups, as far as possible from this one and each other without overlap. K-means (Bishop, 1995) is just one possibility for grouping; a more sophisticated alternative is the mean shift.

Mean shift finds different cluster centres (or pdf modes) by always moving the centre of the cluster to the centre of mass. This is re-evaluated at each step from the new clusters and involves a gradient descent algorithm. This is the method that we used to evaluate the groupings below.

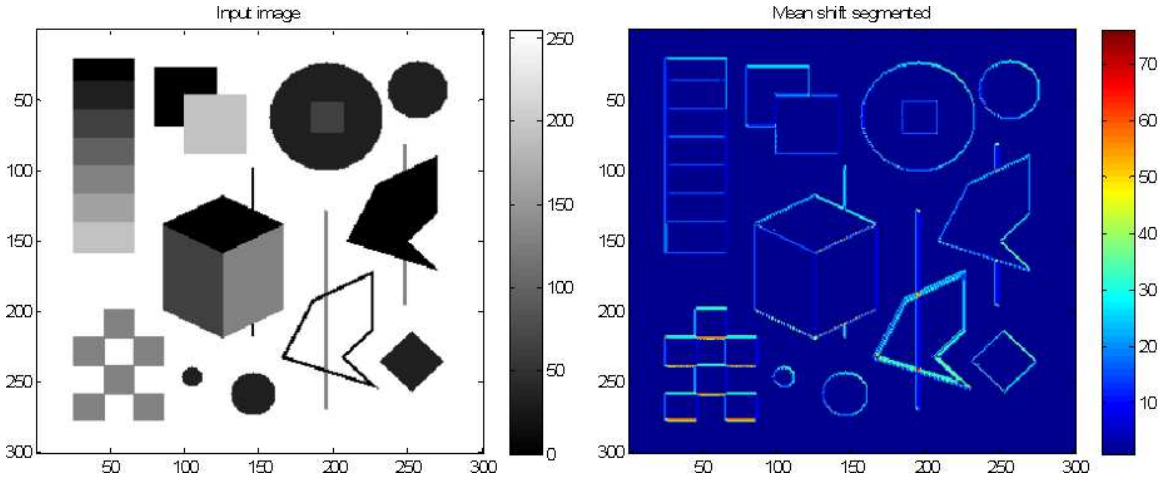


Figure 4.5: Pointwise grouping on multiple scales using the mean-shift algorithm. Each pixel is associated with fu, fv and LP in consecutive three scales. Grouping is performed using the mean-shift algorithm (Comaniciu and Meer, 2002).

To test the power of the Riesz components, each pixel is associated with its (fu, fv) and LP on multiple scales. One particular example is the S. Smith test image (Smith, 1992) shown in Figure 4.5. On the left hand side the original image is shown; this image was designed specifically to combine many of the challenges of feature

detection into a single non-textured image; the noise-free case is shown here. The points to note are: varying edge contrast (left), and various steps and lines/ridges along oriented structures in the rest of the image. The grouping (detected pdf modes) using the mean-shift algorithm (Comaniciu and Meer, 2002) produces a clustering that is appropriate for a phase congruence image. This is shown in Figure 4.5 right hand side.

In computer vision a common practice to look at the local pixel connectivity is by the so called local binary patterns (Wang and He, 1990). This is obtained by some form of similarity measure between the centre of the patch of interest and its neighbourhood. To investigate whether this proves to be sensible for genuine Brodatz texture samples we have evaluated the differences w.r.t. intensities. This local correlation is shown in Figure 4.6 that is calculated as the scalar product of the differences from the centre pixel demonstrated on the Brodatz figures as:

$$Corr(x, N) = \prod (x - N_i) \tag{4.4}$$

where $Corr(x, N)$ denotes the correlation of pixel x with its neighbours N_i and $\prod$ means scalar product. The higher its value the lower the similarity to its neighbours is. That is why it acts as a feature detector.

Overall, the conclusion of the grouping experimentation is that the (fu, fv) pair along with LP is a useful feature descriptor on multiple scales as was illustrated by the experimental result presented in Figure 4.5.

4.2.3 Riesz weighted LP

The use of the Riesz components as a source of structural information, even if they are not individually rotationally invariant, was shown experimentally in the previous section. We have also illustrated through those examples that the LE is not useful in

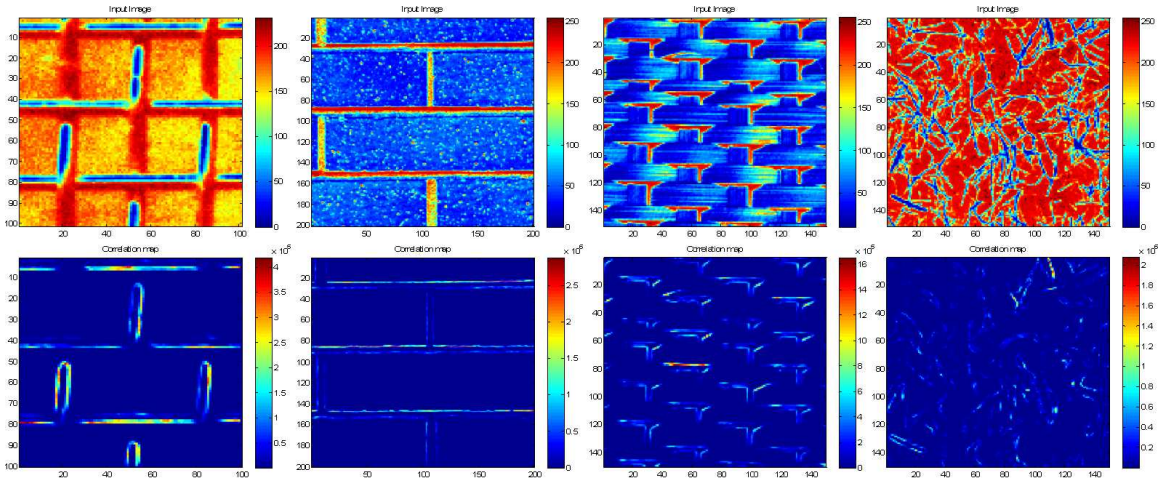


Figure 4.6: Local correlation on the Brodatz samples. The first row shows the original input image, the second row shows the correlation maps. The columns are as follows: Brodatz 1, Brodatz 26, Brodatz 65 and Brodatz 110. The local correlation was measured as the scalar product of the differences between the centre and the eight neighbours, normalized by the square root of the sum of squares.

most of the cases to quantify featureness. This leads to the formulation of the Riesz weighted LP in the first instance, then the Riesz–weighted phase congruence in the second. The Riesz weighted phase is defined as:

$$\mathrm{LP}_{\mathrm{RIESZ}}(\overline{x}) = \frac{\sum_{s_k=1}^{N} \left(\max_k |\mathbf{R}_1(s_k)| + \max_k |\mathbf{R}_2(s_k)|\right) \cdot \phi\left(\overline{x}, s_k\right)}{\sum_{s_k=1}^{N} \left(\max_k |\mathbf{R}_1(s_k)| + \max_k |\mathbf{R}_2(s_k)|\right)} \tag{4.5}$$

where $\overline{x}$ is position, s denotes scale, k denotes the running parameter of the scale and the rest is as defined in the previous chapter, i.e. the $\mathbf{R}$s are the Riesz components and ϕ is phase.

We find that the Riesz–weighted PC gives better results than the energy weighted PC. This is illustrated in Figure 4.7 and 4.8, 4.9. These figures compare the energy weighting and the Riesz weighting, as marked in the image title. The first is the leopard image while in the second and third image there are two texture mixtures from the Randen and Husoy texture set (Randen and Husoy, 1999). Note the differences in dynamic range between the energy and Riesz weighted LP, in favour of the latter. Also, note that the details present in the difference image are due to high frequency

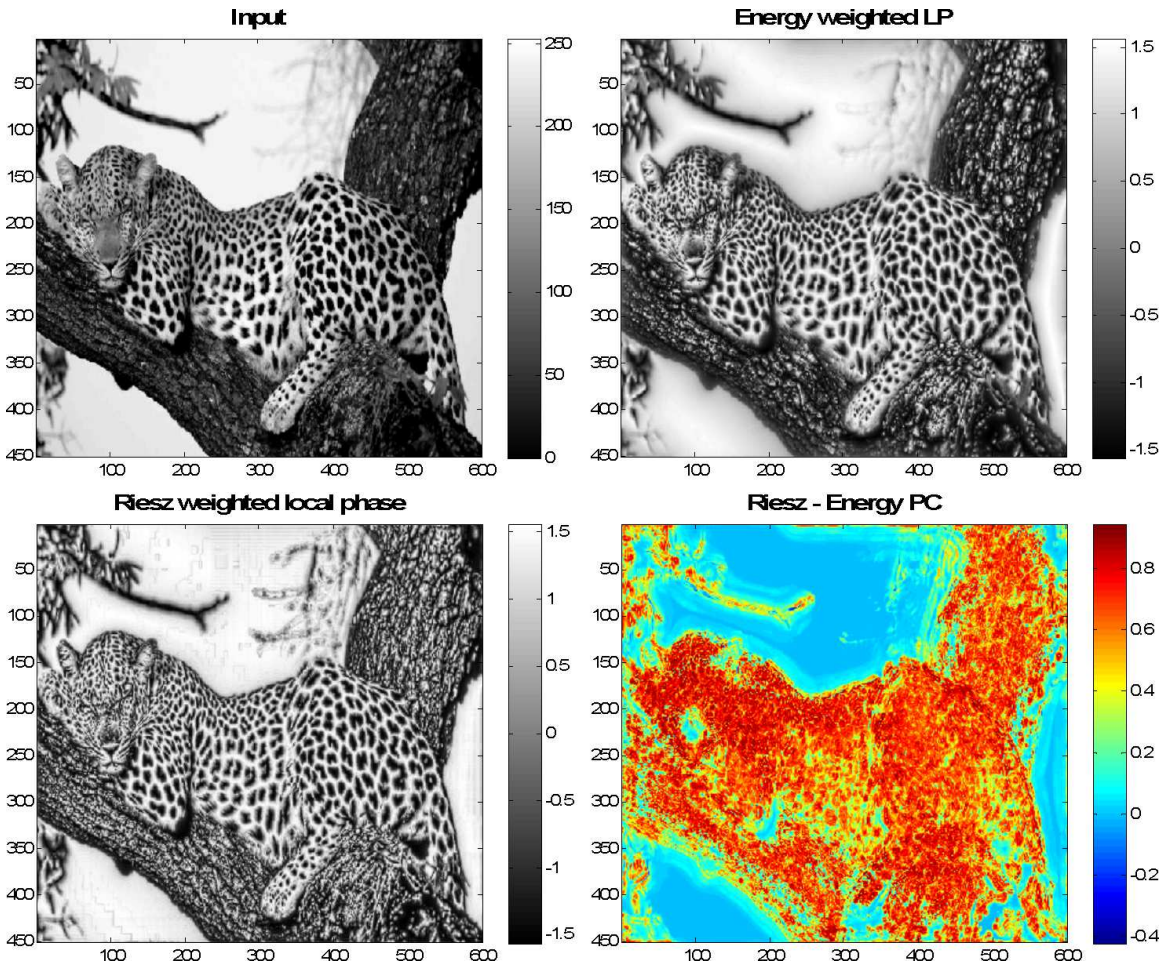


Figure 4.7: Leopard: Riesz weighting is better than the energy weighting. The images from left to right, top-bottom are as follows: original input image, the energy weighted phases, the Riesz weighted images, and the difference between the corresponding PCs. Note the significant textural information of the rims of the texture patches that are present in the Riesz weighted phase, in red on the final difference image.

textural changes present in the Riesz weighted PC as is demonstrated in the remainder of this chapter.

Figure 4.7 shows both the Riesz and the energy weighted LP. To be able to compare the performance of the two measures the PC is calculated in both cases and the difference is shown in the last figure, this has a value between 0 and 1. This image shows that the Riesz weighted PC captures most of the inside textural part of the patches instead of the edges, which is why the red parts appear in the final image.

The changes in the Figure 4.8 are more specific. The Riesz weight captured

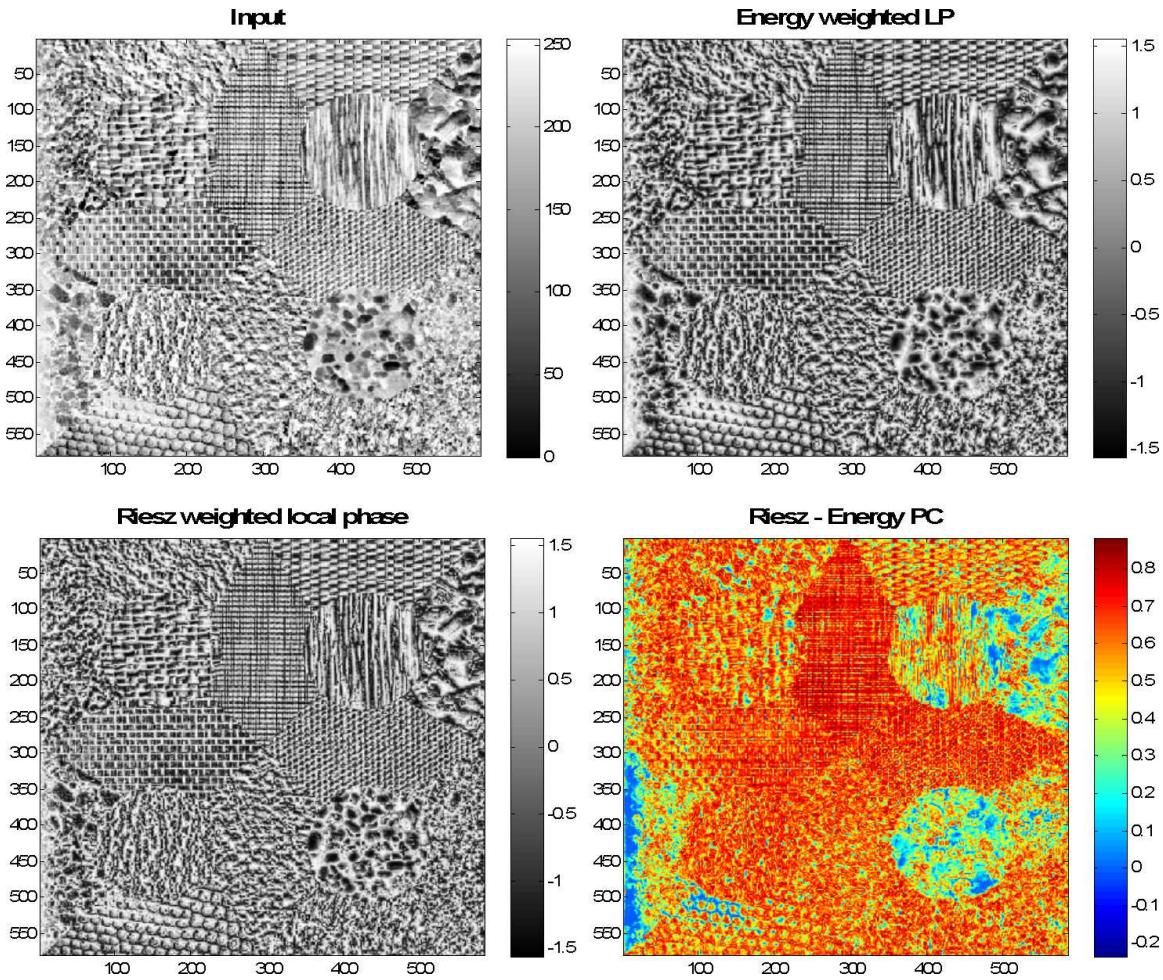


Figure 4.8: Randen and Husoy texture mixture I: Riesz weighting is better than the energy weighting. The figures in order from left to right and top to bottom: original input image, the energy weighted phases, the Riesz weighted images, and the relative difference between the Riesz and energy weighting PCs that correspond to it. Note the differences of high frequency textural information that is present in the former, but not the latter one.

textures are clearer and this is shown in several parts of the image both of which are represented by blue in the final difference image. These are the parts that are more blurred for the energy weighted local phase map on the right hand side image. Similar observations can be made in Figure 4.9, where the clearer regions are towards the right hand side and middle of the image.

Figures 4.10, 4.13, 4.14, and 4.15 also compares the energy weighted LP with the Riesz weighted phase, but now with a range of different types of image. From

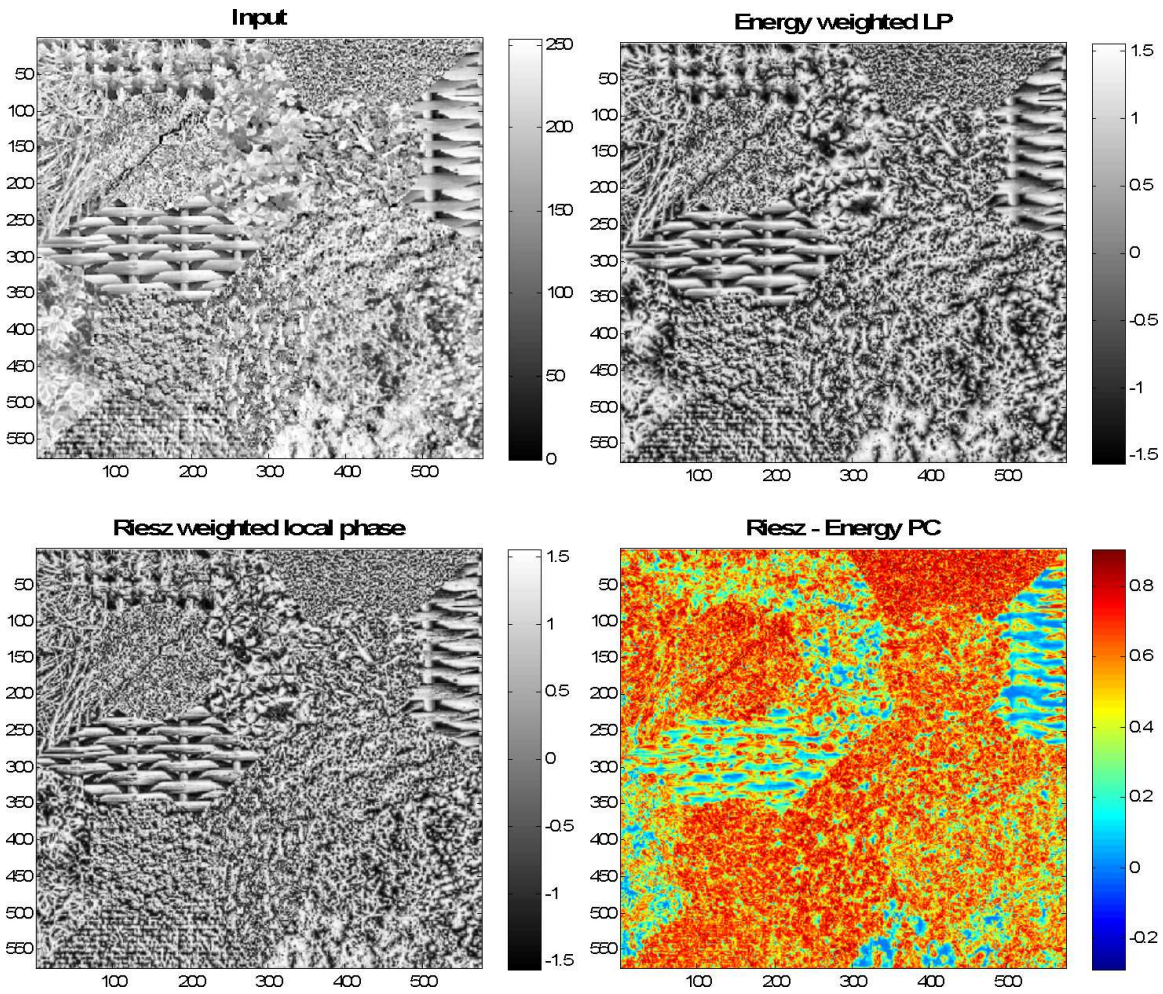


Figure 4.9: Randen and Husoy texture mixture II: Riesz weighting is better than the energy weighting. The images from left to right and top to bottom are as follows: original input image, the energy weighted phases, the Riesz weighted phase, and the relative difference between the latter two. Note that the positive values maximized at the red colour in the final difference image are textural information that is present in the Riesz weighted PC but not the energy weighted one.

left to right and top down the images are as follows: the input intensity, the energy weighted phase, the Riesz phase, finally the difference between the Riesz weighted PC and the energy weighted PC. Observe, that there are differences in the detail between the energy and Riesz weighted PC. These are exactly the high frequency texture descriptors that we are interested in. Note, for example the extra details that appear on the checks, feathers and background in the Lena image. The fact that the Riesz weighted PC contains extra texture instead of noise is demonstrated by the

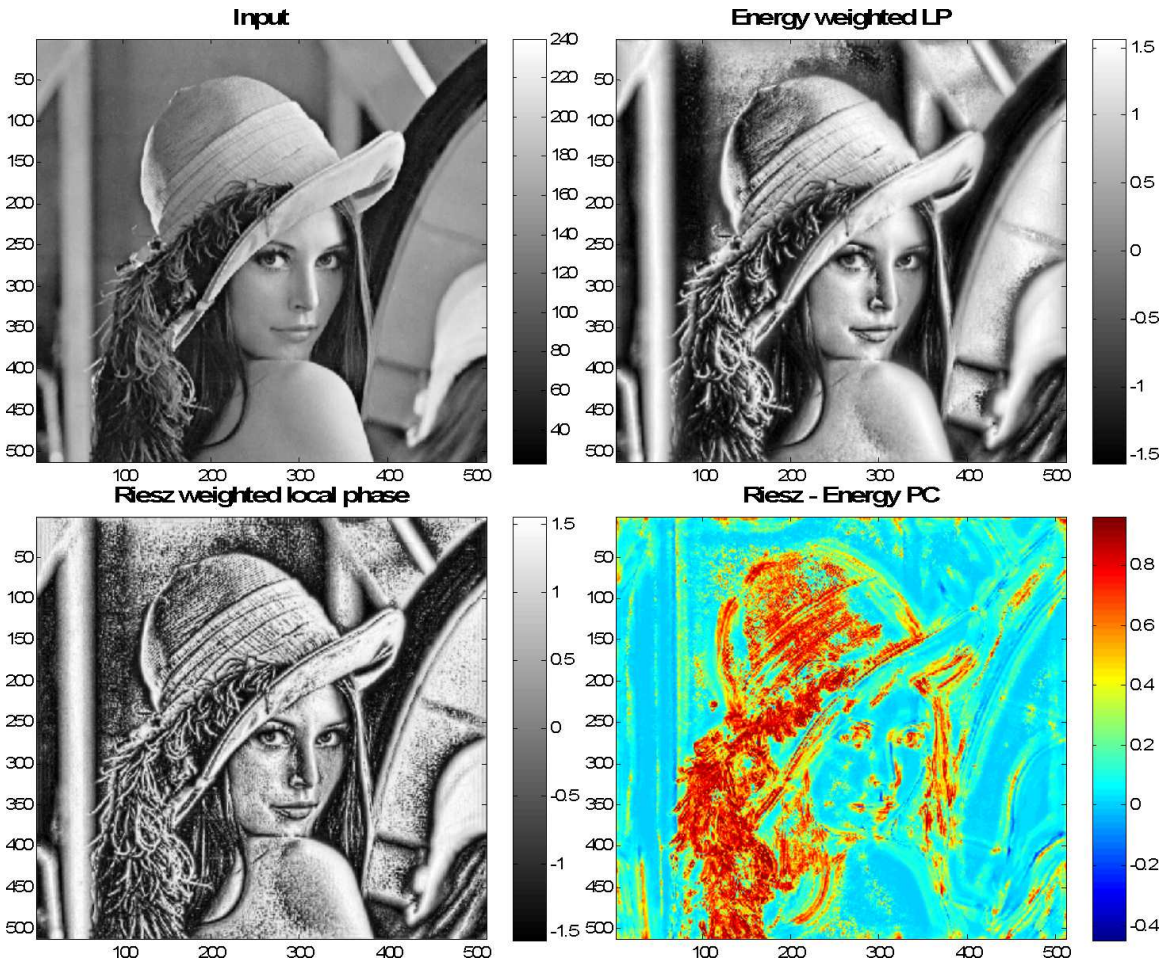


Figure 4.10: Differences between the Riesz weighted and energy weighted PC for the Lena image. The images clockwise are as follows: original input image, energy weighted LP, relative differences between the corresponding PC maps and the Riesz weighted LP.

magnification of the part of the feather and the background in Figures 4.11-4.12.

Figures 4.11 demonstrates that the extra parts of the Riesz weighted LP contains feature instead of noise. Follow the arrows and discover that the energy weighted PC completely loses the curvilinear structure at the arrow, while the Riesz weighted ones preserve it and this can be seen from the difference image itself. Figure 4.12 shows that both edges are detected by the Riesz weighted component as opposed to the energy weighted component. This is also visible on the difference image.

The Riesz weighted PC also has the advantage of better detecting repeating struc-

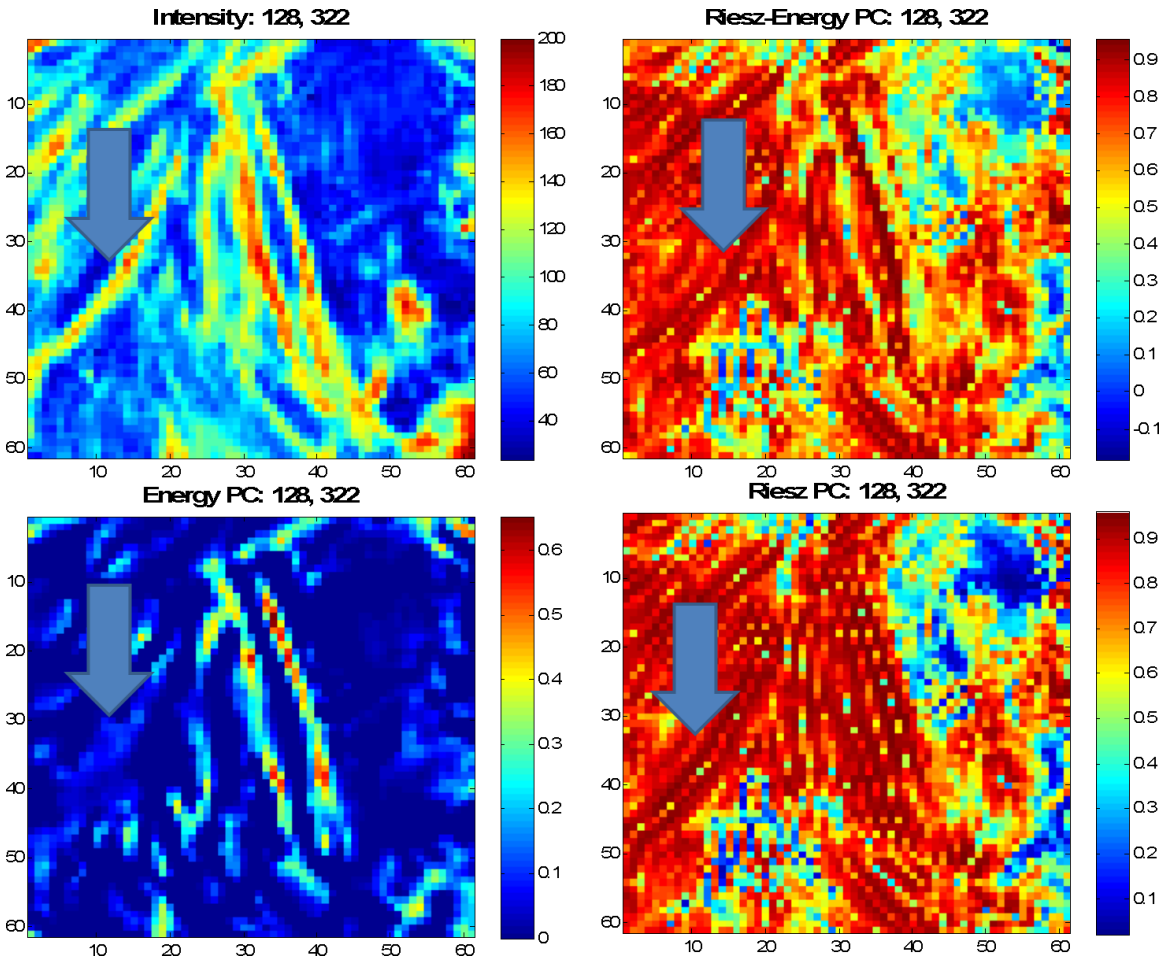


Figure 4.11: Structures on Lena's hat: differences between the Riesz weighted and energy weighted PC at the location specified in the title. The images clockwise are as follows: original input image centred on the given feather position, Riesz - energy weighted PC, energy and Riesz weighted PC. The arrow points to a curvilinear structure of the hat, which is also visible on the intensity image, however as opposed to the Riesz weighted PC the energy weighted one does not detect it.

tures, for example, the bricks that compose the chimney of the house, a magnification of which is shown in Figure 4.14. Although little may be said about the zebra image of the correctness (which here is defined by the detection of the stripes on the zebra) of the Riesz weighted PC as opposed to the energy weighted one, we contend however that it is quantified properly by the difference image as is shown in Figure 4.16. This figure contains a part of the zebra's leg which as we observed from the input image has a weaker contrast to the rest of the body. This has a strong influence on the en-

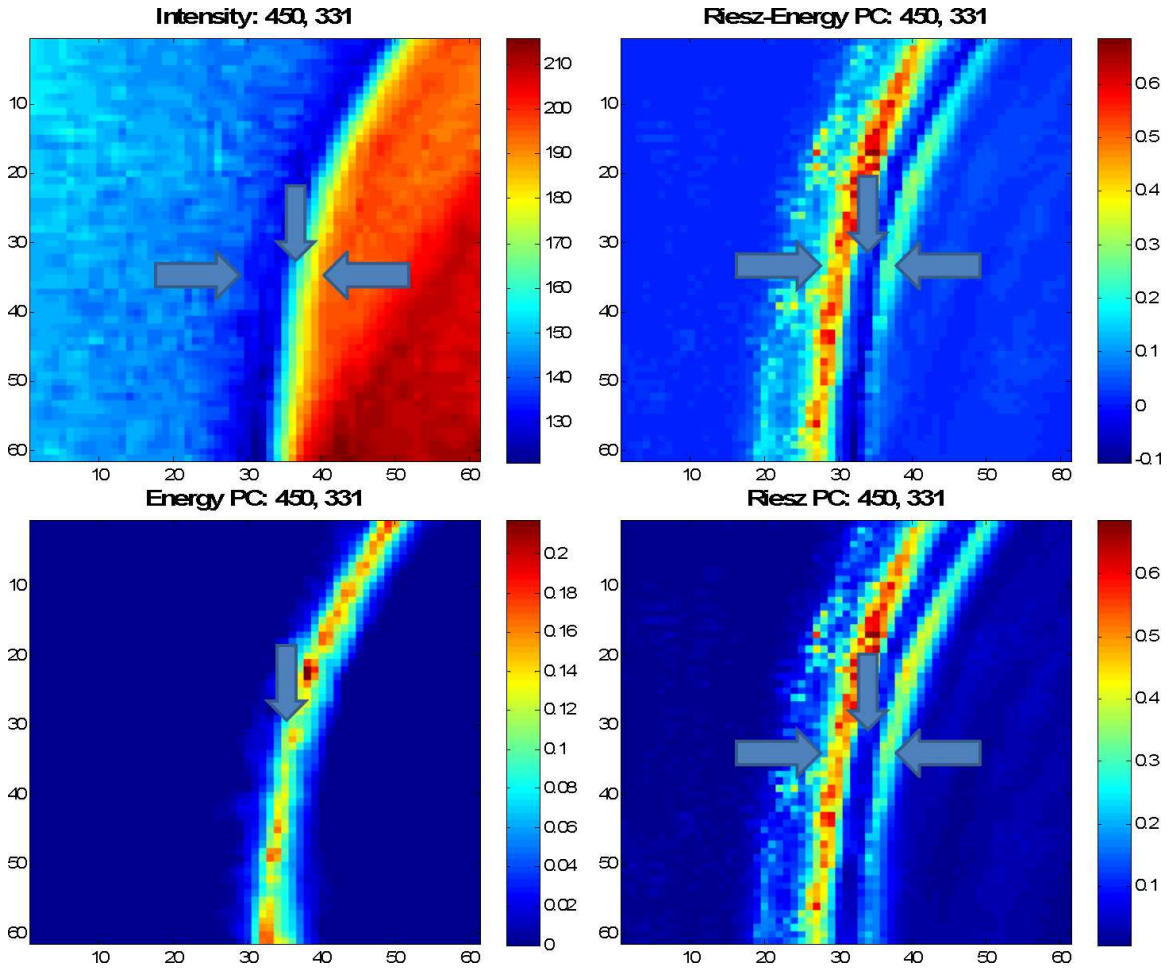


Figure 4.12: Curved edge on the RHS of Lena: differences between the Riesz weighted and energy weighted PC at the location specified in the title. The images clockwise are as follows: original input image centred on the given position, Riesz - energy weighted PC, Riesz and energy weighted PC. The energy weighted PC has a smaller magnitude and only detects a single edge as compared to the Riesz weighted PC. The arrows positioned in the same direction on the different images show the same features, the curvilinear edge/ridge positions.

ergy weighted PC resulting in low or extinct edge information. This is demonstrated on the above figure where the arrows point to the original stripes.

The extra piece of information detected is part of the structural description and not noise. This is demonstrated by all PC differences shown above. Red shows the structures that are present in the Riesz weighted PC but not the energy one. These final difference images and the magnified parts show that as well as texture information as in Lena's hat and the bricks in the house, ideal edges and ridges are

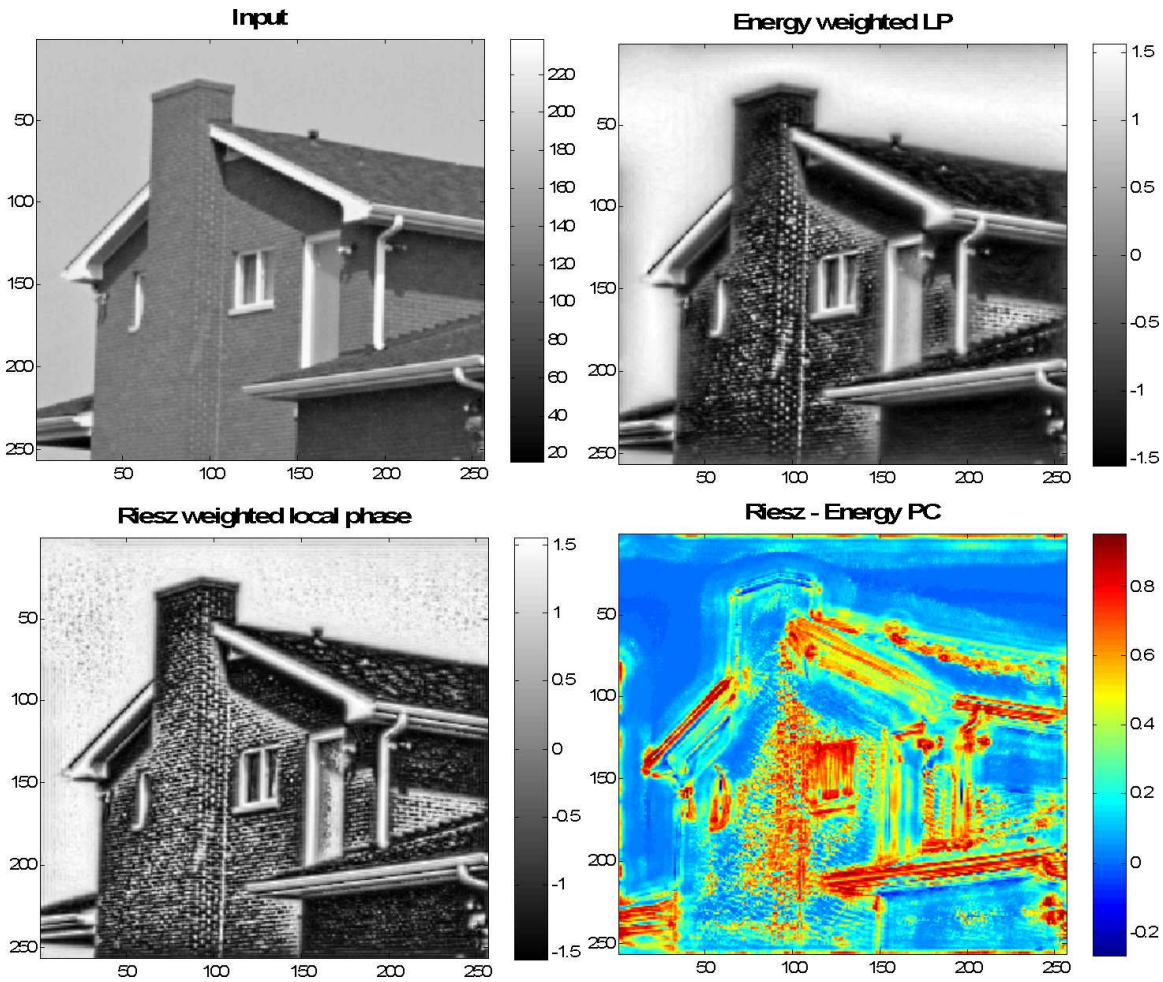


Figure 4.13: Differences between the Riesz weighted and energy weighted PC for the house image. The images clockwise are as follows: original input image, energy weighted LP, Riesz-energy weighted PC, and Riesz weighted LP. Note that although is hard to perceive with the untrained eyes the differences between the two weighted LP maps, however, these clearly show up on the difference between the corresponding PC maps in favour of the Riesz weighted methodology.

also revealed in all three images (see Figure 4.12).

Figure 4.14 shows the extra piece of information present in the Riesz weighted PC as opposed to the energy weighted one. The arrows point to the regular structures that are detected on the texture of the chimney, however, not on the energy weighted version. The arrows are in correspondence, therefore the exact same features may be compared on the intensity and the Riesz weighted images. As already mentioned, unfortunately not much can be seen on the energy weighted PC in this scenario.

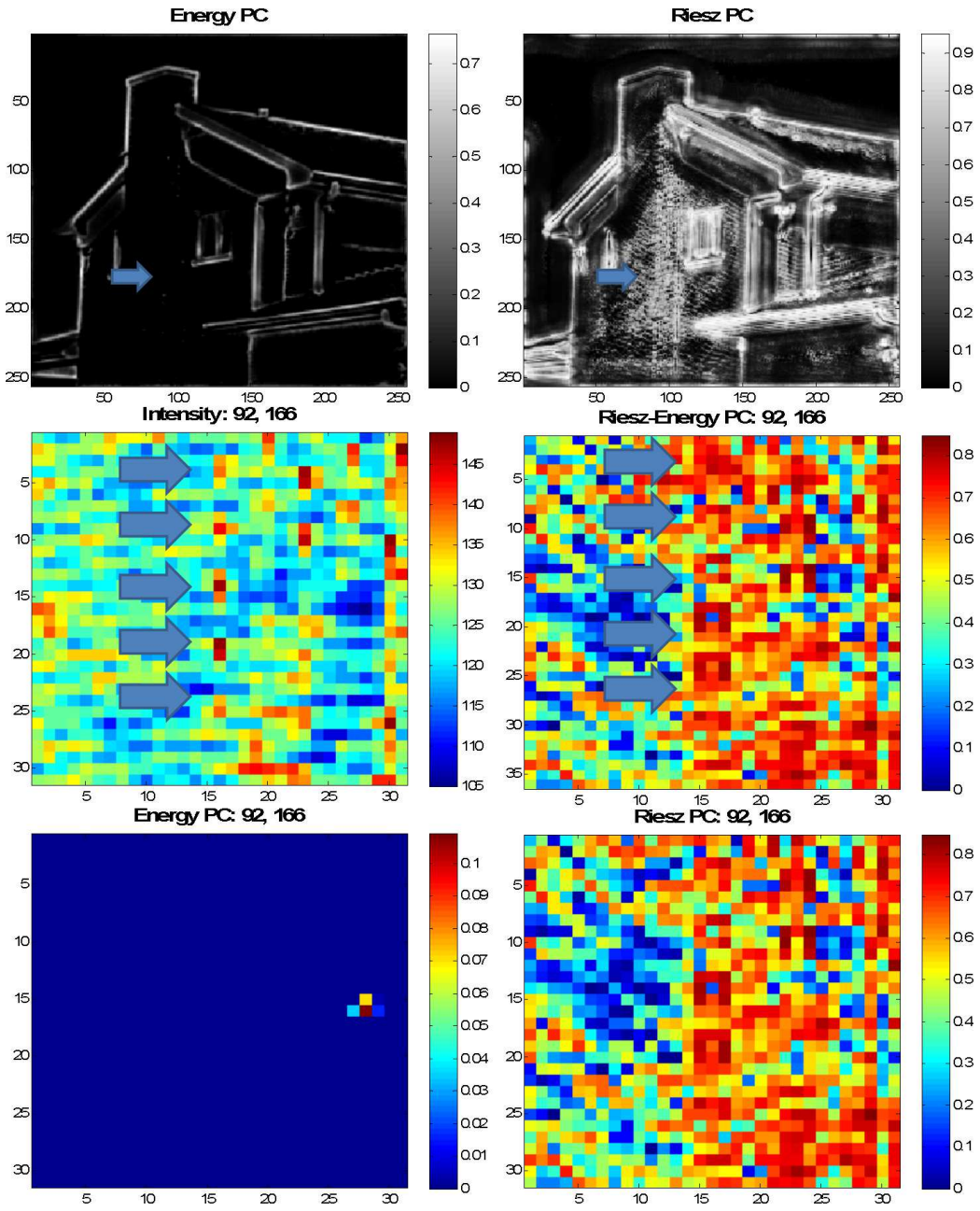


Figure 4.14: House bricks: differences between the Riesz weighted and energy weighted PC in detecting the bricks of the chimney. The images are as follows: first row shows the energy and Riesz weighted PC, the second row magnifies the region around the position given in the image title for the intensity and differences between the corresponding PC maps, finally, the third row shows the magnified regions on the PC maps shown in the first row.

A further demonstration that we detect features instead of noise may be seen in Figure 4.16, which shows part of the leg of the zebra. The arrows point to the regions

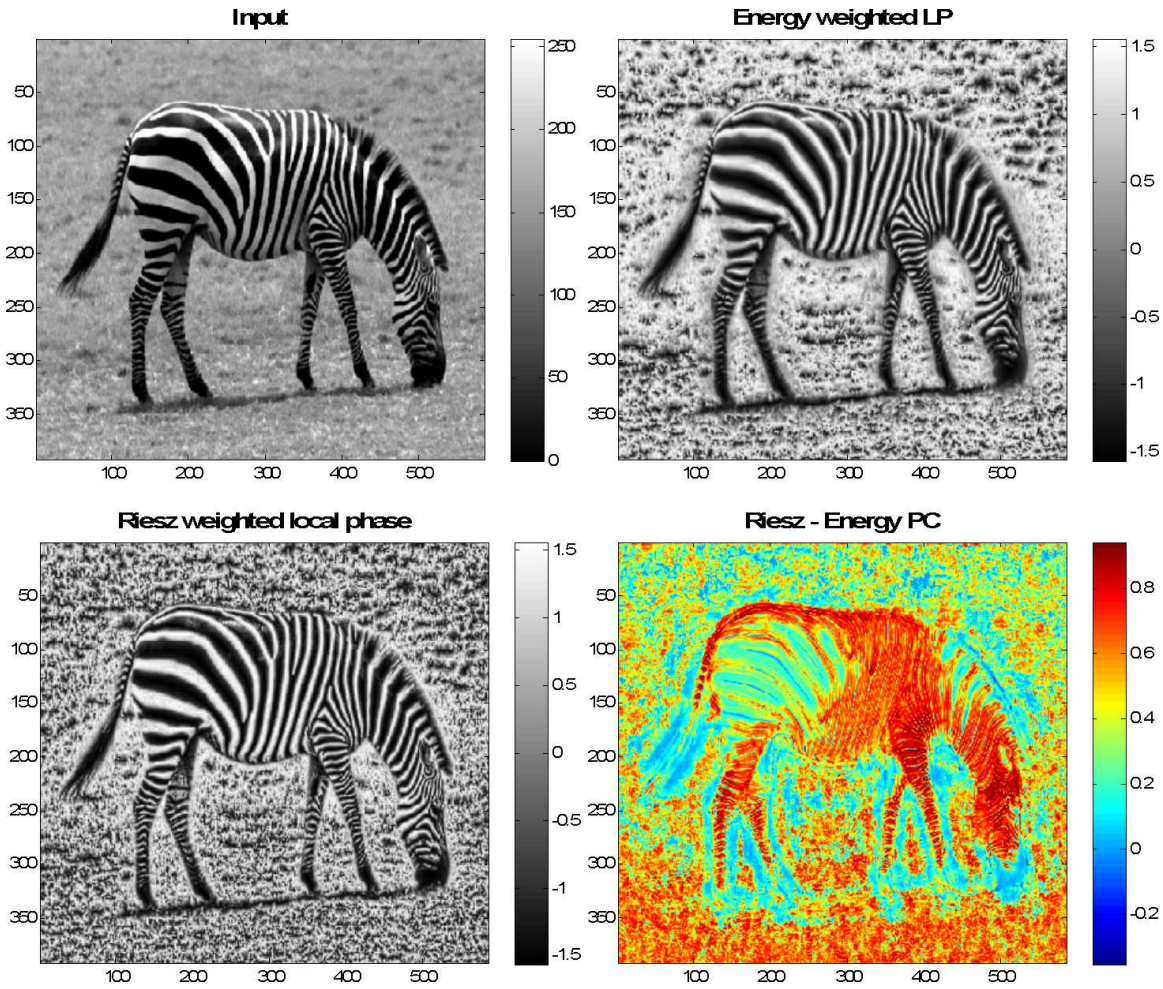


Figure 4.15: Differences between the Riesz weighted and energy weighted PC for the zebra image. The images clockwise are as follows: original input image, energy weighted LP, the differences between the corresponding Riesz and energy PC maps, and the Riesz weighted LP.

where a feature should be present due to the observations made on the intensity images. We can observe that the features that are present on the Riesz weighted PC and the energy weighted PC are very different in consistency. The latter one is less consistent and occasionally disrupted while the Riesz weighted signal gives a fuller description of the stripes that are originally present in the images.

One application of the difference between the energy weighted and the Riesz weighted LP is automatic initialisation of liver segmentation. There are several semi-automatic segmentations that benefit from seed points to initiate them. One such

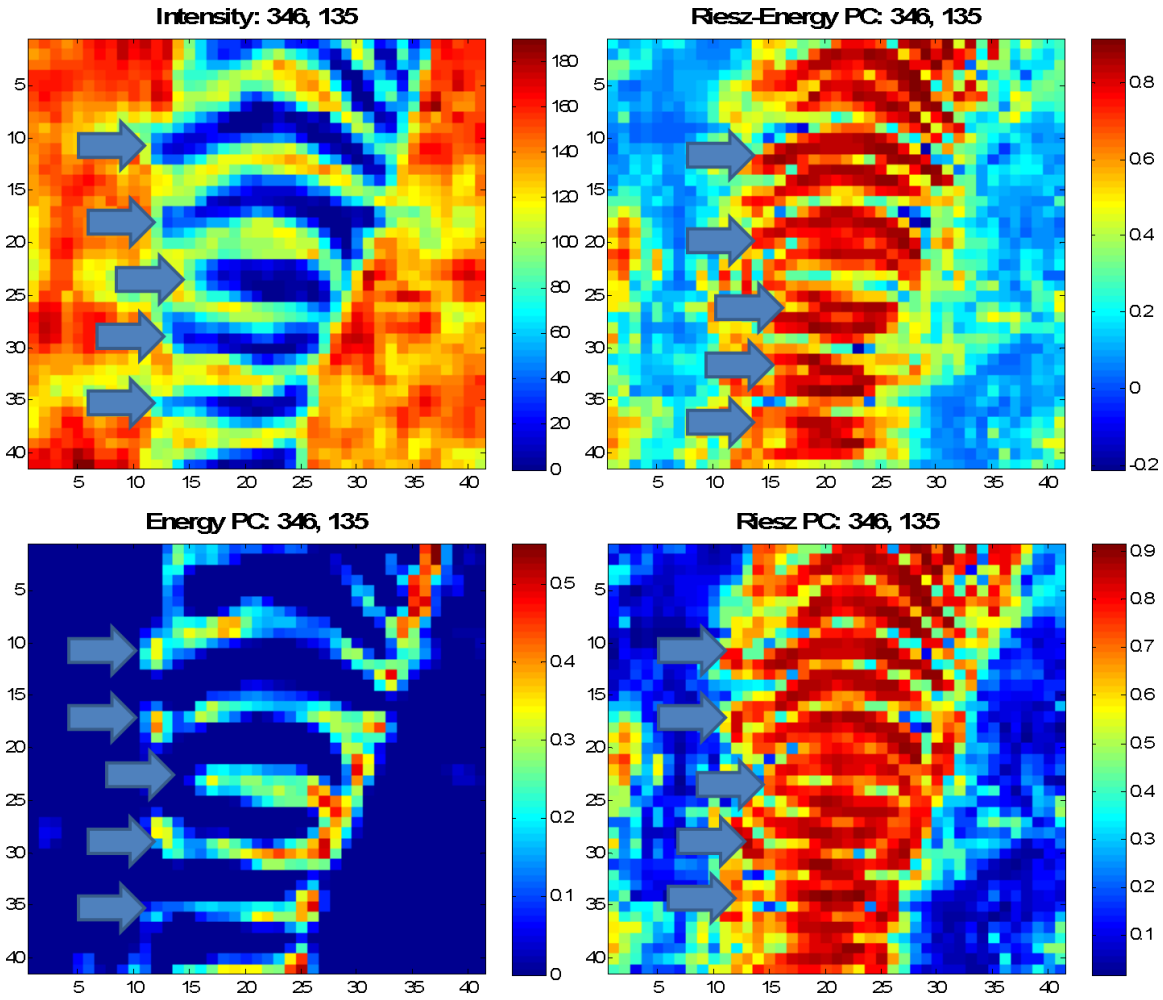


Figure 4.16: Stripes on the zebra's legs: differences between the Riesz weighted and energy weighted PC for the region centred around the position given in the title of the image. The images clockwise are as follows: intensity image, difference between the energy and Riesz weighted PC, Riesz weighted PC, and energy weighted PC. Note that the stripes on the legs from the energy weighted PC are detected to a less extent due to reduced contrast on the intensity image. As opposed to this the Riesz weighted PC detects the boundary of the stripes as is annotated by the arrows.

algorithm is the level-sets (Caselles et al., 1997). However, seeding these algorithms is a time costly process and biased by the person performing the task. Taking the difference between the energy and the Riesz infinity norm weighted LP produces consistent results.

The difference between the energy and Riesz weighted LP maps may be used to seed the level set algorithm for intensity based liver segmentation. It was ob-

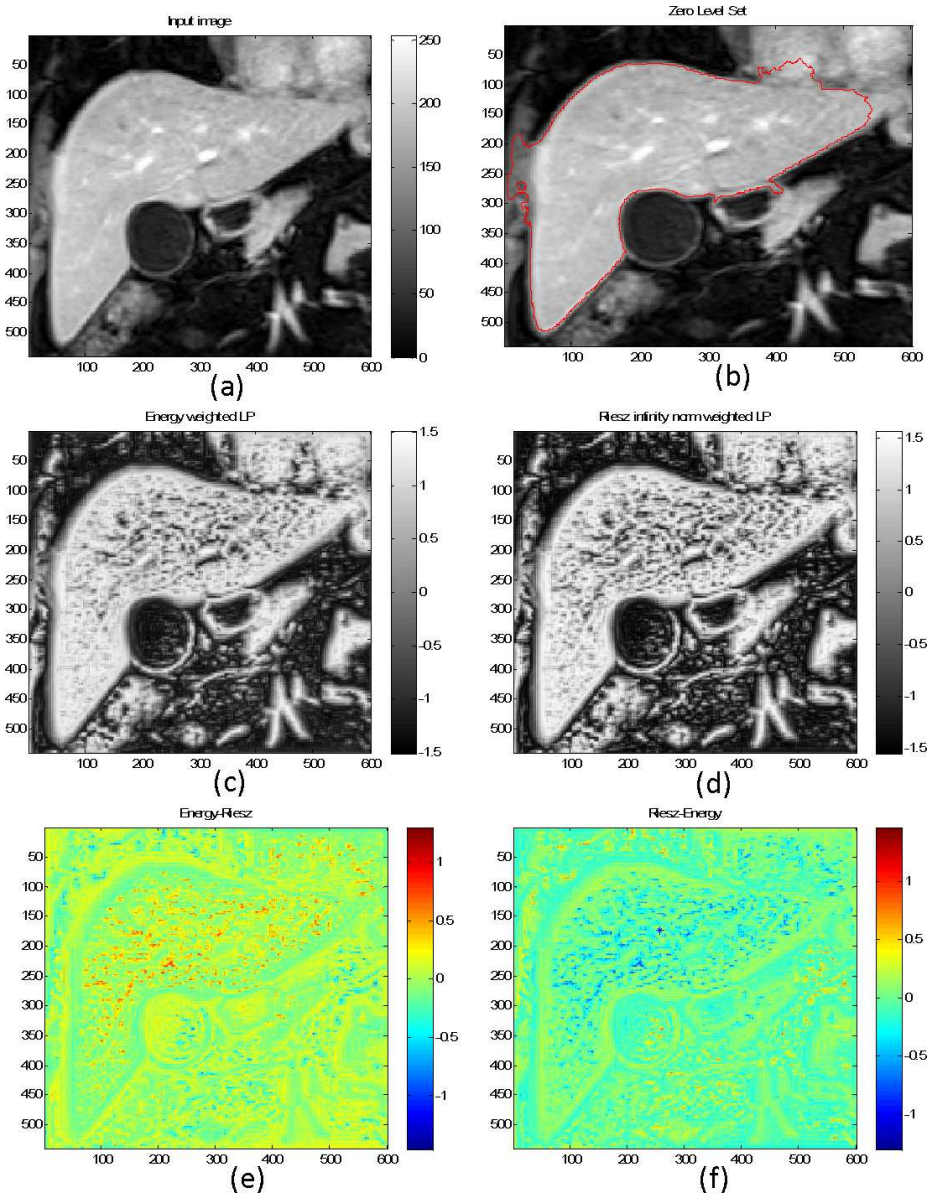


Figure 4.17: Automatic seeding of level-set for liver segmentation. The first row from left to right shows the original image and the segmented image. Note that there are overflows of this segmentation, but is not because of the seeding, and that the level-set segmentation is not a major component of this application. The second row from left to right shows energy weighted local phase (LP) and the Riesz weighted LP. Note the increased structures available for the Riesz weighted LP. The final row shows the difference image between the energy weighted and the Riesz weighted map on the left and the other way round on the right. The black star on the last figure shows the selected seed point with the proposed methodology.

served that the difference between the energy and Riesz weighted LP is consistent with respect to features that characterize the liver. One particular example of the

segmentation is shown in Figure 4.17. The first row shows the input image and the segmented image. The overflow in the segmentation is not because of a wrongly selected seed point. Rather, it is a feature of the algorithm itself, the development of which is not the central question addressed in this part of the thesis. The energy weighted and the Riesz weighted LP maps are shown in the next row. The final row shows the difference images, the left is the difference between the energy weighted and the Riesz weighted LP maps while the left one is taking the difference the other way round. Our algorithm used the second difference formulation, and the selected seed point is indicated by the black star on this map. In conclusion, it produces a seed point that is more than enough for a level-set segmentation algorithm. A similar method may be devised to seed the algorithm with whole areas; but that is left for future work.

Another important aspect of the figure is the comparison between the energy and Riesz weighted PC in the middle, which demonstrates that the Riesz weighted method detects more structure than its energy weighted counter part.

4.2.4 Texture measures using the Riesz weight

As a qualitative example of texture we compared the local entropy maps on different Riesz weighted local phase feature maps. These results are shown in Figure 4.19-4.22. The images illustrate the 3x3, 5x5, and 7x7 entropy maps of the Riesz weighted local phase maps. We first present the calculations, then the motivation behind the choice of the particular methods, and finally conclude by discussing the results.

The entropies are calculated based on probability density functions estimated with the Non Parametric Windows method (NP win) (Joshi, 2007). The Shannon entropy was calculated in local windows as:

$$H = -\sum (\overline{x}) \ \log p (\overline{x}) \tag{4.6}$$

where the H denotes the entropy at location summed over $\overline{x}$, and p denotes the corresponding pdf.

One interpretation of the pdf is as the mathematical, continuous version of a histogram, which is a discrete approximation to the pdf, albeit one that is a very good approximation if there are sufficient samples (law of large numbers).

That is, it counts the frequencies of occurrence of intensity values in the image disregarding any spatial organization. In a histogram, values of the parameter (eg pixel intensity) that are "close" are counted into the same bin. One must be careful in transforming histograms into pdf-s taking in account the bin width and sample size. Given that all spatial order is thrown away very different images may have the same pdf. We can imagine an arbitrary image and the image that is composed of the exact same pixels but in which the pixels are randomly redistributed, these will have the same histogram (overall, it is exactly not for texture at all!), though of course they will appear very different as images. To avoid this limitation, methods that take in account close neighbourhoods have been devised, such as the kernel density method (Parzen, 1962).

One particular problem with kernel density approximation is speed. In fact the requirements on the pdf estimators are illustrated in Figure 4.18. The Figure contains the pdf of the eye image situated on the top left hand side with different approaches. The right hand side of the Figure down shows a table that situates the available methods based on the requirements, which include accuracy and speed. We observe that histogram and non-optimal kernel density estimates have high computational efficiency but provide poor accuracy. Optimal kernel density methods in turn are computationally inefficient, however, provide accurate pdf estimate. The NP windows method is a relatively new technique that is able to reproduce the optimal kernel density estimator's performance by taking in account pixel ordering, critical sampling, and the fact that images are bandlimited. NP windows provides a fast and

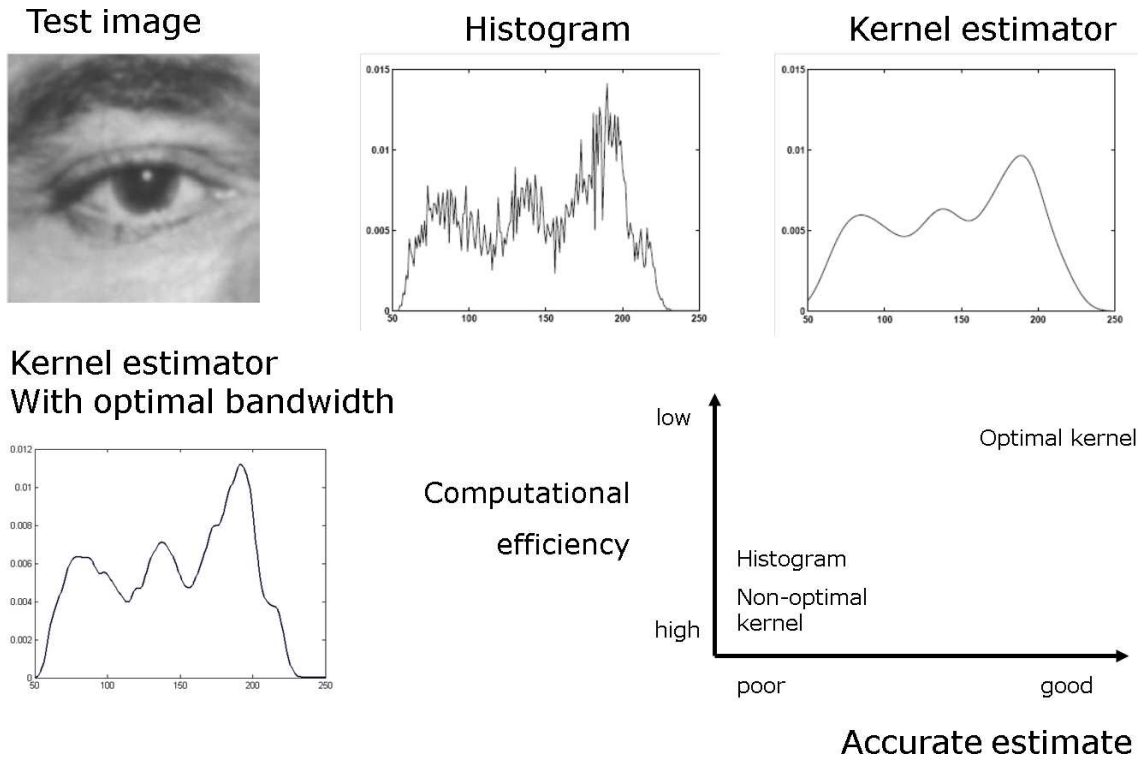


Figure 4.18: PDF estimation using different methodologies. This figure was adapted from one of Sir Michael Brady's slides.

accurate estimate with a relatively small sample size. Moreover, it is completely data driven, without the assumption of the Gaussian model that underlies the kernel density estimator. This is why it is desired for textures. Based on these considerations, we used the NP windows methods. Figure 4.19-4.20 and 4.21-4.22 first column shows that we achieve a reasonable differentiation with the local entropy maps between the different textures on different scales. There are two scales used in this context, the base image is composed of several filter scales (usually 3-4), and there is the size of the window in which the pdf is estimated and the entropy is calculated. We observe in the first column of Figure 4.19-4.20 that by increasing the window size (NP win scale) we are still able to differentiate between the different Randen and Husoy texture mixture components. A similar observation can be made about the texture of the bricks in the next two images Figure 4.21-4.22. The second column in both cases pose a completely different challenge. We are into identifying edges and ridges and

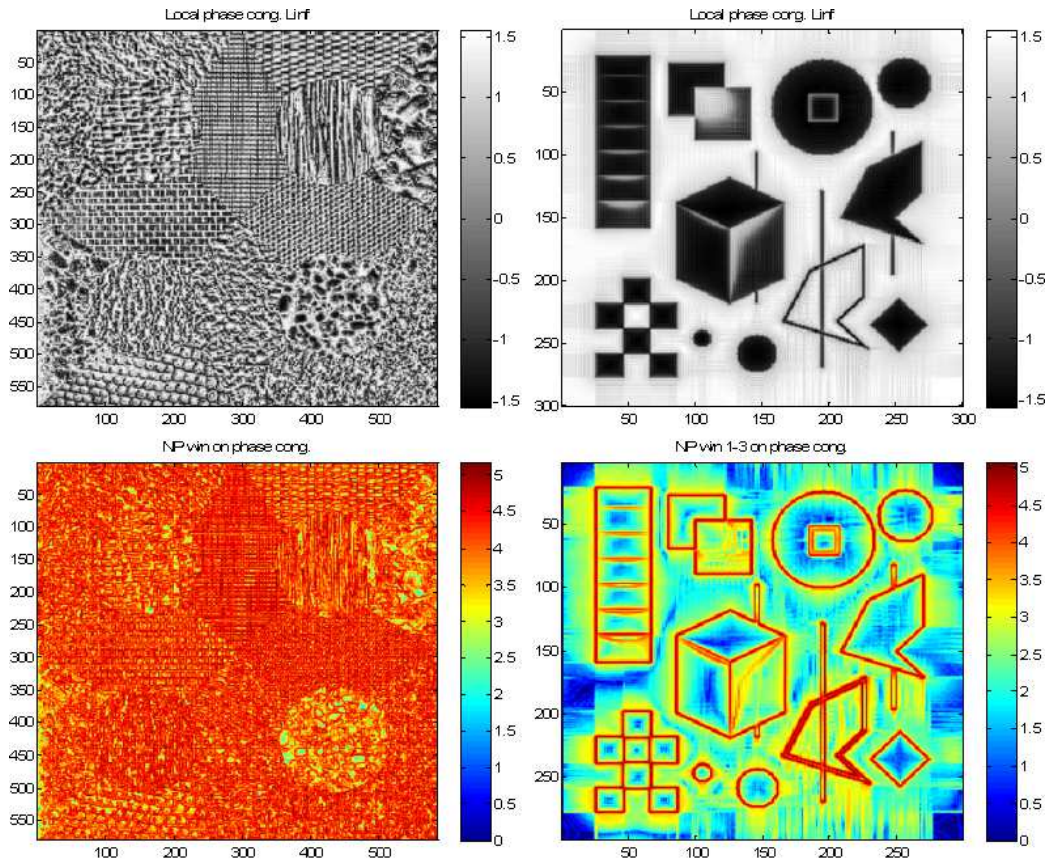


Figure 4.19: Local entropy maps of Riesz weighted local phase values. The rows are as follows: first the weighted LP, then the 3 x 3 NP win entropy map is shown for two test images.

it is necessary that these get wider in size once the scale of the NP windows pdf increases.

As a result, both the use of the Riesz weighted phase maps and NP windows are suggestive of feature energy measures that qualitatively give good separability between different textures.

4.2.5 Riesz weighted PC

The usefulness of the Riesz weighted method is hard to quantify without a measure that is similar to phase congruency. Phase congruency means that the phase values align at feature points, including edges and ridges, and only at such points. The

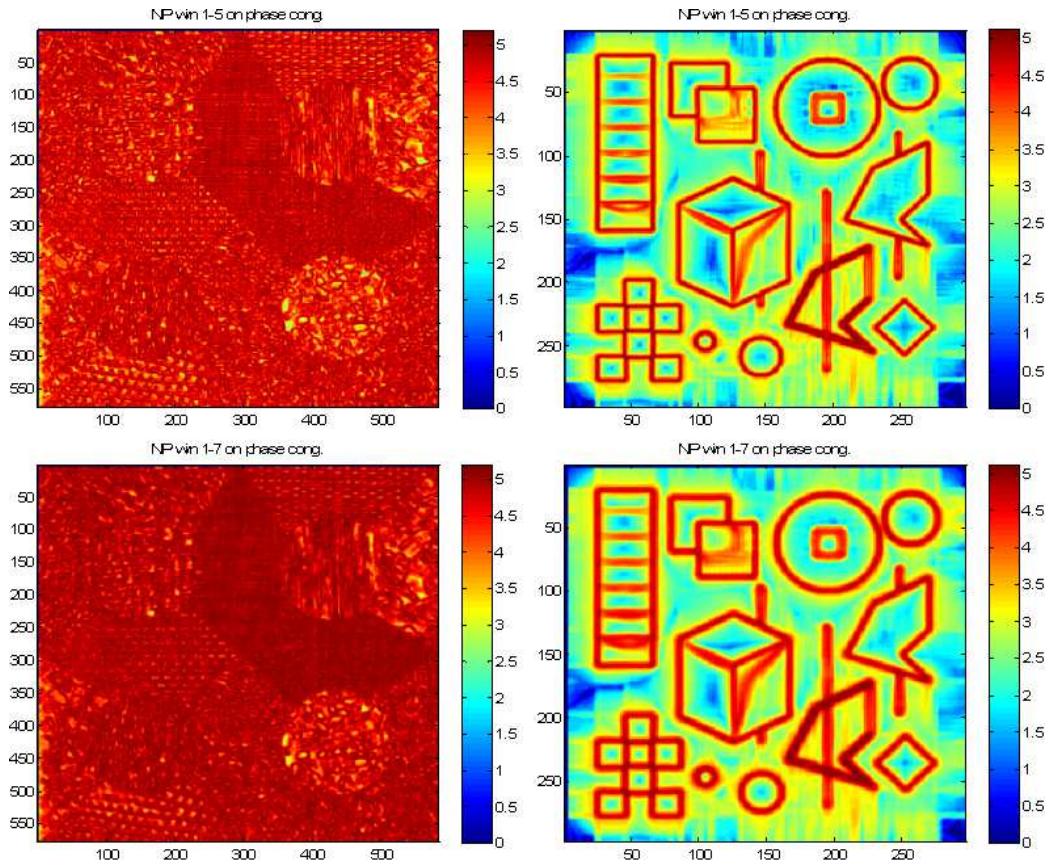


Figure 4.20: Local entropy maps of Riesz weighted local phase values (continuation of the earlier figure). The rows are as follows: first row is the 5 x 5 NP win entropy map and finally the fourth row signifies the 7 x 7 entropy map.

original idea of phase congruency was suggested in (Morrone et al., 1986). It means that the Fourier series expansion components align at the same phase value for edges and ridges and therefore evenness and oddness may be differentiated as was discussed earlier. A very good description of this theme is in (Gosta and Knutsson, 1995). As already described, its extension to images was not until the energy model was published (Morrone and Owens, 1987) (Morrone and Burr, 1988).

The energy model states that we have maximal phase congruency where the energy attains a maximum. In other words, the energy of the analytic signal signifies where the features are. This led to the energy weighted phase measurements to determine phase congruence by (Kovesi, 1999). The original model was formulated using phasors

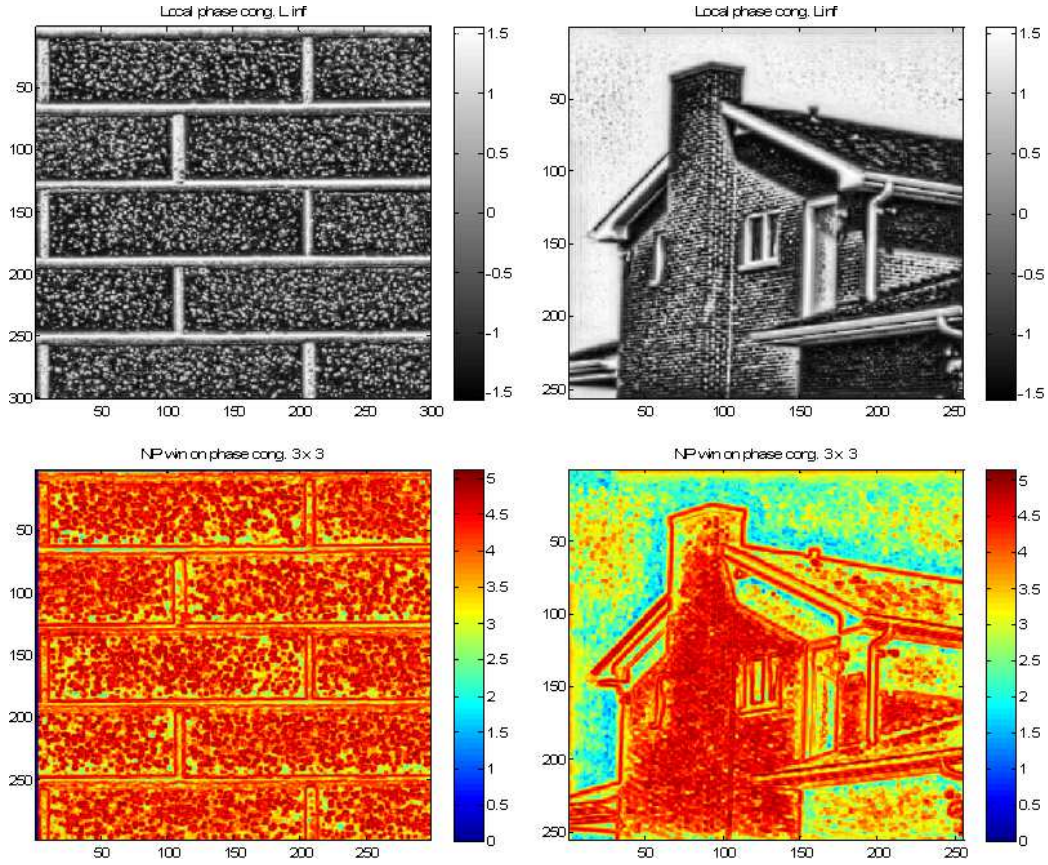


Figure 4.21: Local entropy maps of Riesz weighted local phase values. The rows are as follows: first the weighted LP, then the 3 x 3 NP win entropy map.

as:

$$PC(x) = \max_{\overline{\phi} \in [0,2\pi]} \frac{\sum_n A_n \cos(\phi_n(x) - \overline{\phi}(x))}{\sum_n A_n} \tag{4.7}$$

where x is a particular image location, A_n is amplitude and ϕ and $\overline{\phi}$ denotes the phase and average phase. In practice, there are several modifications that need to be made to this theoretical model; but those considerations are omitted from this short survey. Instead we return to the practical formulation of our problem.

Especially, when looking at quantification of texture versus noise, to prove its advantages, we need to find a way to quantify featureness using the Riesz weighted phase values. To do this, we modified the original phase congruence implementation provided on-line by Kovesi to contain the Riesz weights instead of energy (Kovesi,

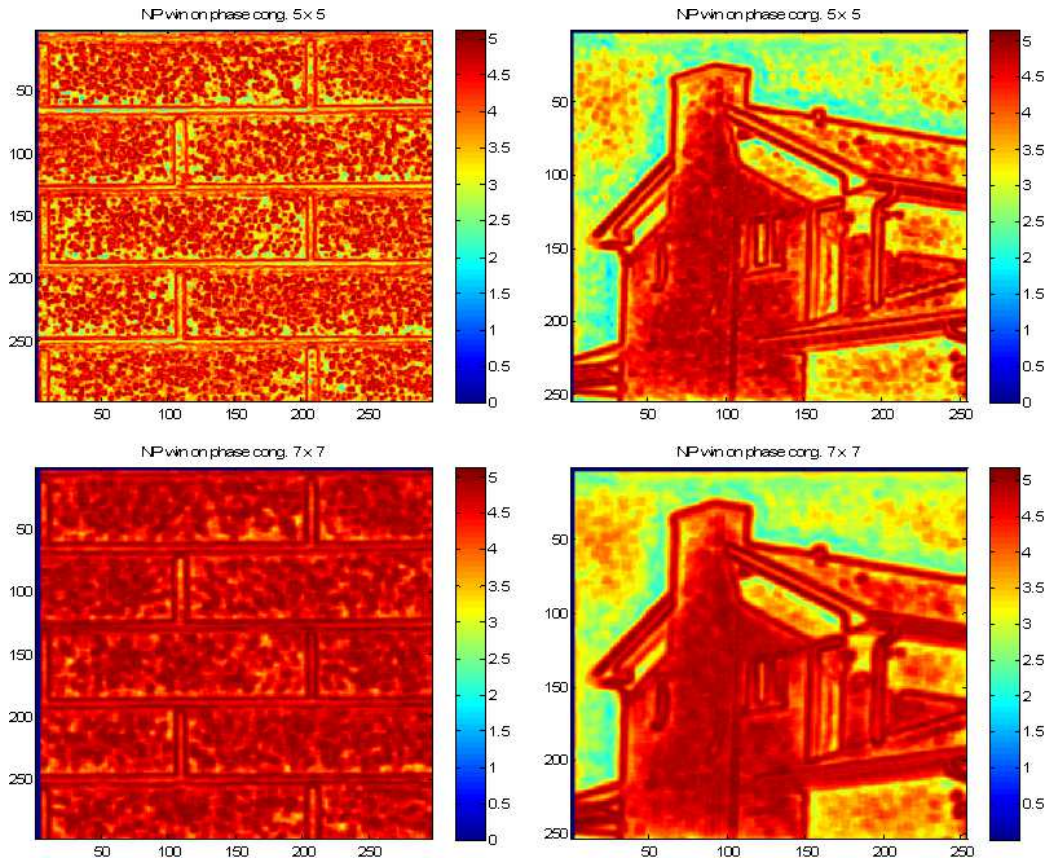


Figure 4.22: Local entropy maps of Riesz weighted local phase values (continuation of the previous figure). The rows are as follows: first row 5 x 5 NP win entropy map and finally the fourth row signifies the 7 x 7 NP entropy map.

1999). The new measure is more prone to detect features that are omitted with the energy weighted version. We argue that this is because of the lack of information in LE sometimes as shown earlier in this chapter in Figures 4.1- 4.2.

We compared the phase congruency with energy weighting and Riesz weighting. We illustrate this in Figure 4.23, where we show the energy weighted and the Riesz weighted PC for different stages of liver disease. The rows from top to bottom signify stage two through stage four of liver disease. Note that there is not much difference in the appearance on the energy weighted images, however features are clearly seen to change in the case of the Riesz PC. Additionally, it is somewhat disappointing that the energy weighting method shows similarity for the stage 2 and 4 case, and very little features for the stage 3.

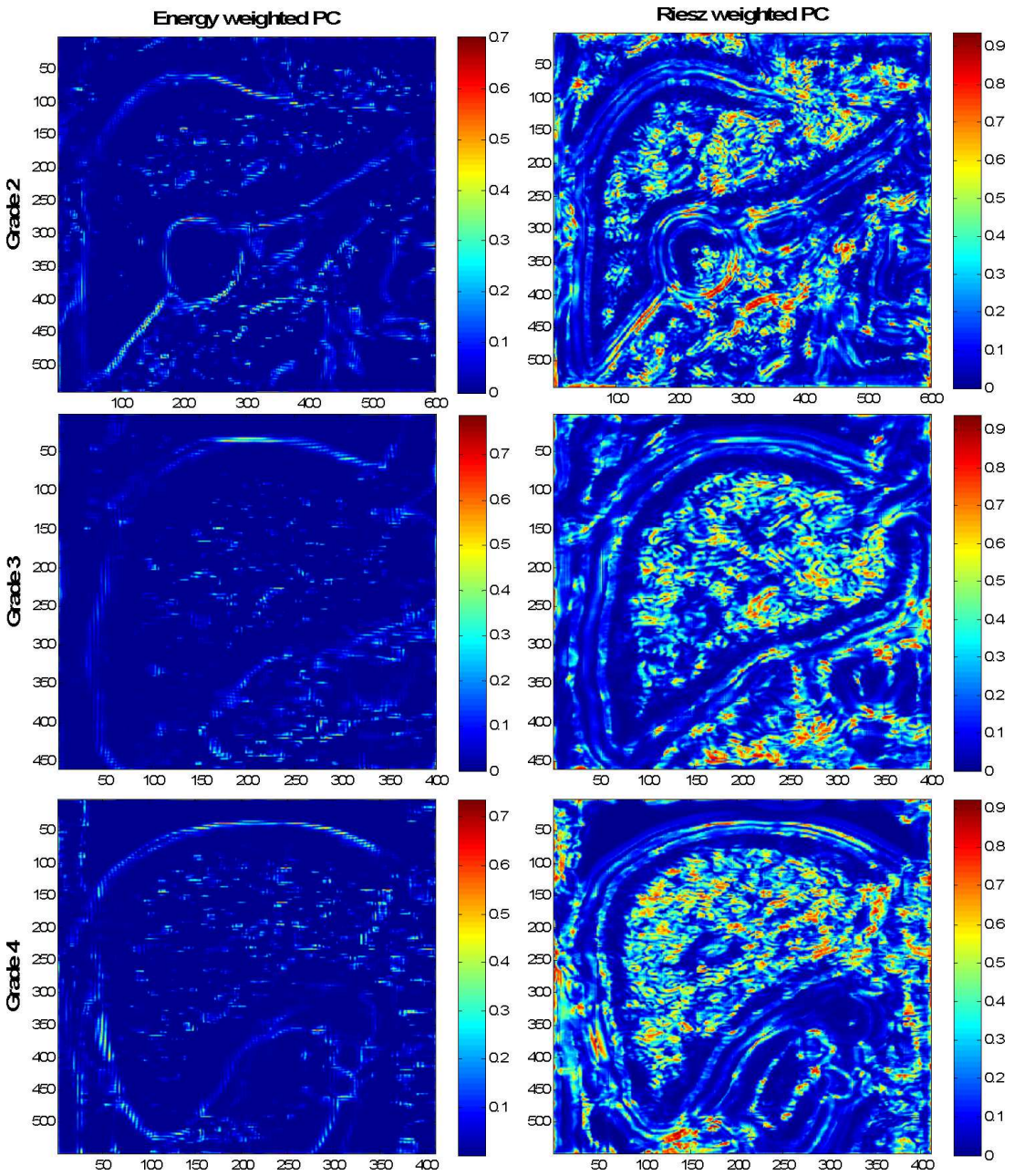


Figure 4.23: Riesz weighted versus energy weighted maps of liver disease. The first column is the original energy weighted phase congruence map (PC), while the second column is the Riesz weighted PC map. The different rows signify different stages of the disease: the first row is stage 2, second row stage 3 and finally stage 4 of liver disease. Note the difference in terms of structural changes that are detected.

The use of the above developed method is further supported by the trend it shows between different stages of liver disease using texture measures. We used fractal dimension as a measure of how self-similar something is, and local entropy

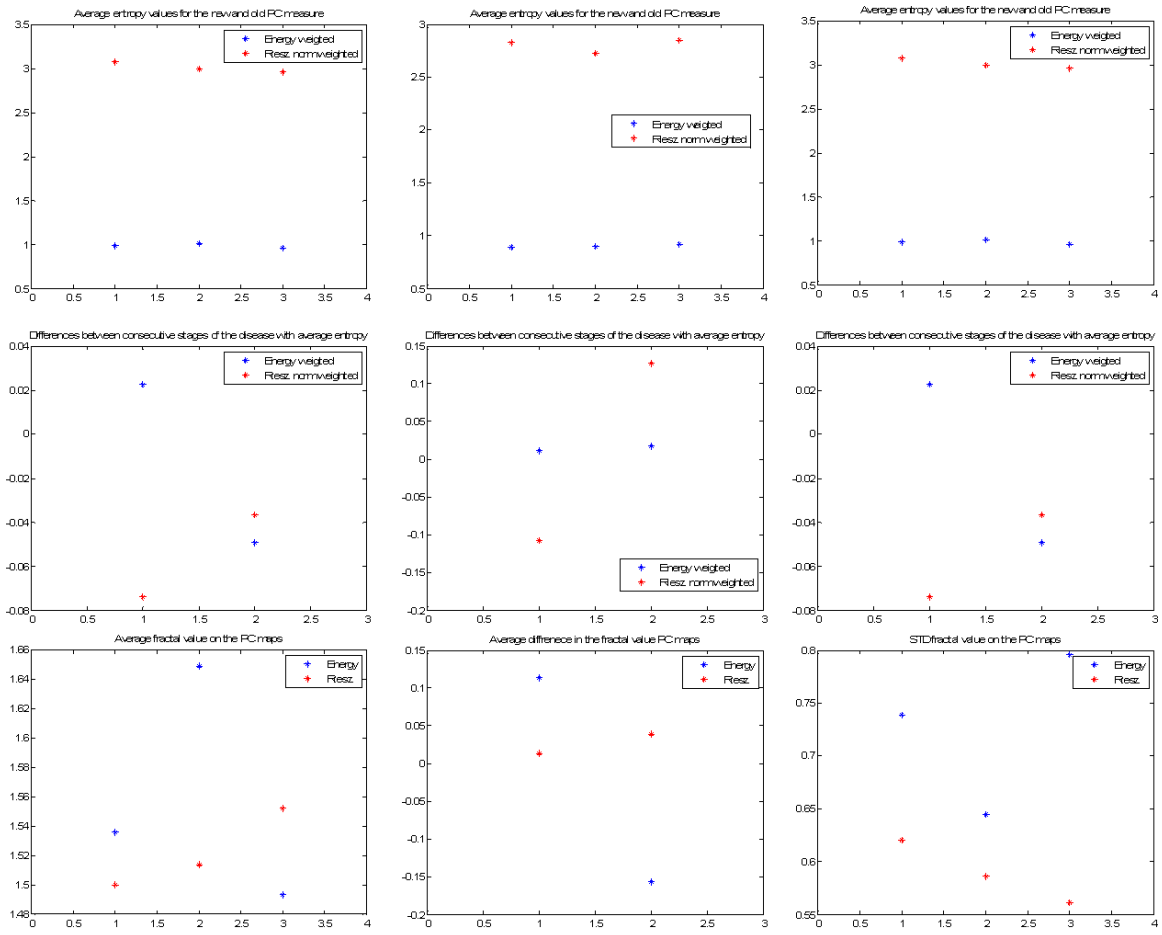


Figure 4.24: Energy versus Riesz weighted maps. The first two rows are measurements from the local entropy maps. The columns here correspond to the different NP win scales with 3, 5 and 7 pixels squared, in this order. The first row shows the average entropy values for the different measures and stages on each of the image. The second row shows the signed difference between the different disease stages for each measure. The last row is a result of fractal measures: the first image shows the fractal measure itself, where we see a definite difference between the two measures. The Riesz weighted one shows a gradual increase. The second image is the same as the measures in the second row, differences between different stages for each measurement. The final image presents the standard deviation.

measurements on the PC maps to look for trends in different stages. We used three scales of local entropy 3x3, 5x5 and 7x7 pixels. The energy and the Riesz weighted PCs were then compared. Both measures decrease to a certain extent; but the measures are not consistent. This is shown in Figure 4.24 first row. We return to texture measures in more details in a later chapter 6.

The second row shows the differences between different stages for each measure. This second row cannot be evaluated because there are flips in the distances that appear with the two measures. This is because sometimes one of the measures does not follow a gradual decrease. The final row shows results calculated from the fractal measure. The first image shows the fractal measure itself, this is consistently increasing with the Riesz weighted measure, but it does not do so with the energy weighted one. Because the grade three fractal measure with the energy weighted measure shows a big step, the second image in the final row cannot be evaluated. The final image is very promising because it shows a smaller standard deviation for the Riesz weighted maps as opposed to the energy maps.

4.3 Fisher Z transform based PC

One way of showing that statistical correlation between consecutive scales leads to sensible texture detection on local phase values both for synthetic and complex liver texture images is considered in this section using the specific measure of Fisher Z transform. It is demonstrated that the PC may be calculated using proper hypothesis testing looking at the cross-scale correlation of the local phase values on consecutive scales. The cross-scale correlations and the hypothesis testing shows encouraging power in picking out structural features. We first show our results on the S. Smith test image (Smith, 1992)(Figure 4.25), then on the Randen and Husoy texture images (Randen and Husoy, 1999)(Figure 4.26), finally, on the liver images themselves (Figure 4.27).

The method proceeds with the calculation of the Fisher Z cross-scale correlation in pixel windows of size 5x5, 3x3 and 2x2, followed by hypothesis testing. The Fisher Z cross-scale correlation windows are only calculated for consecutive scales and the hypothesis testing performed on these. In addition, it is also calculated for the multi-

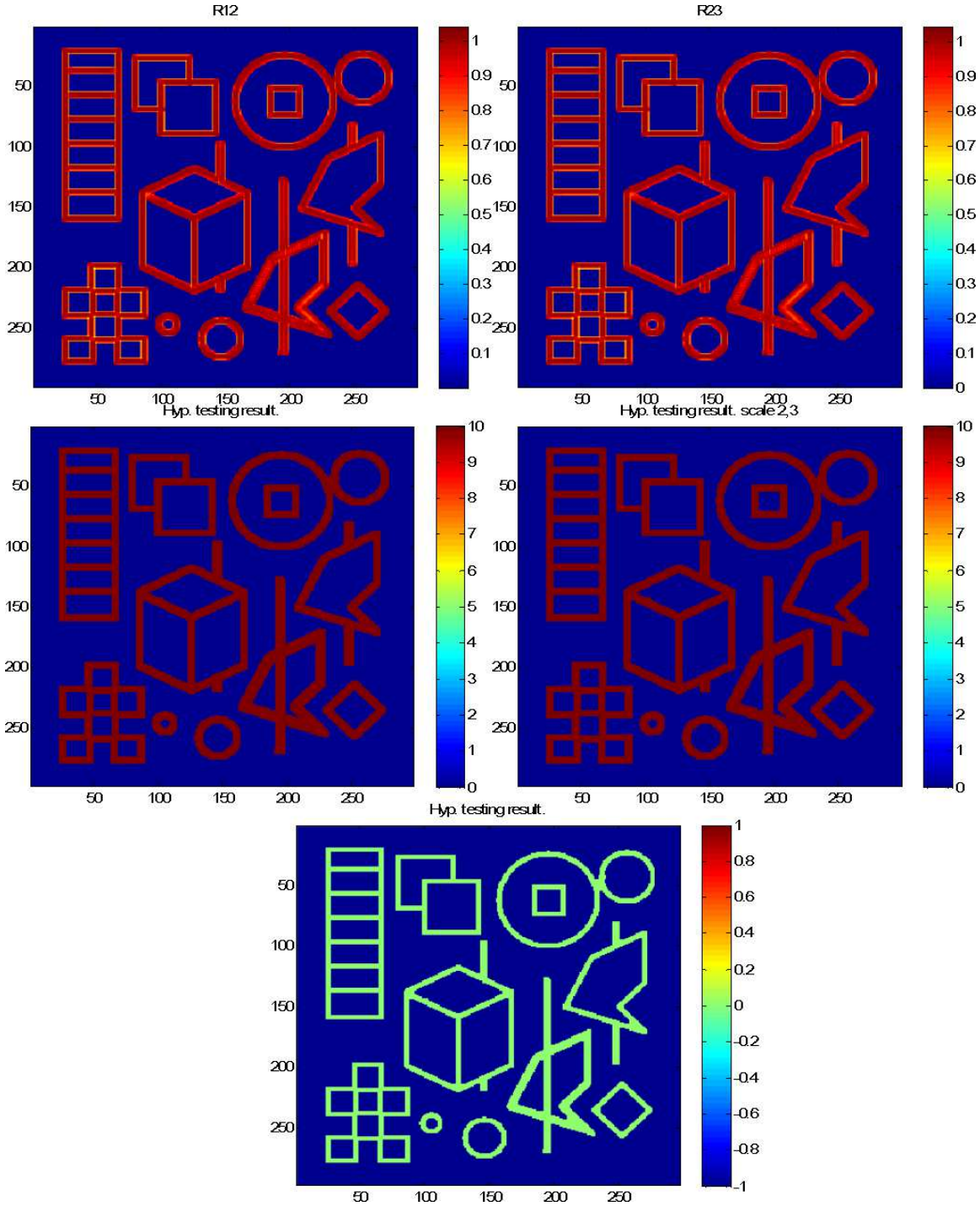


Figure 4.25: Fisher Z transform and hypothesis testing: the first row shows the Fisher cross-scale correlations for scales 1-2 and 2-3; row two shows the corresponding hypothesis testing, final row is the hypothesis test result from the above three scales in total.

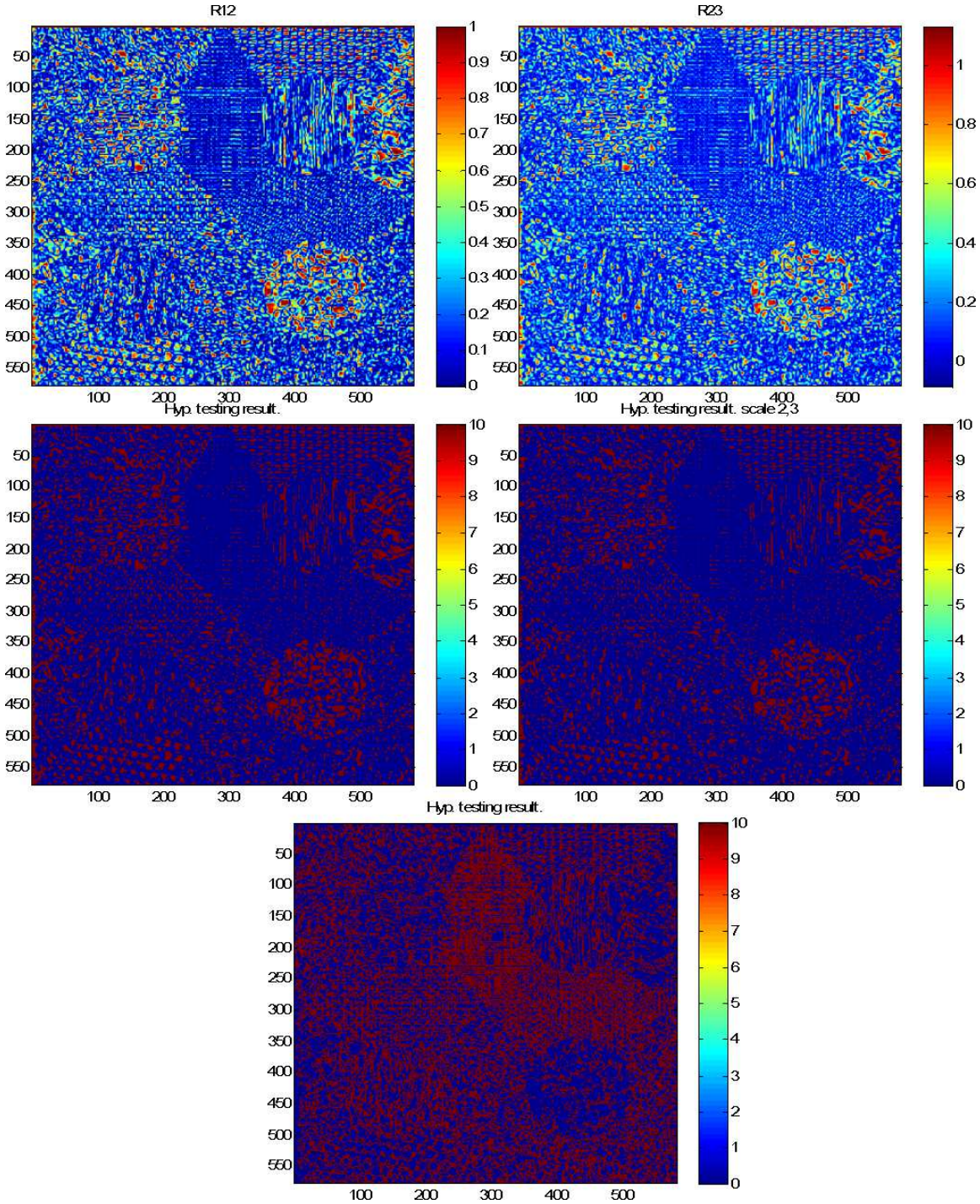


Figure 4.26: Randen and Husoy images: Fisher Z transform and hypothesis testing - first row shows the Fisher cross-scale correlations for scales 1-2 and 2-3; row two shows the corresponding hypothesis testing, the final row shows the hypothesis test result from the above three scales in total.

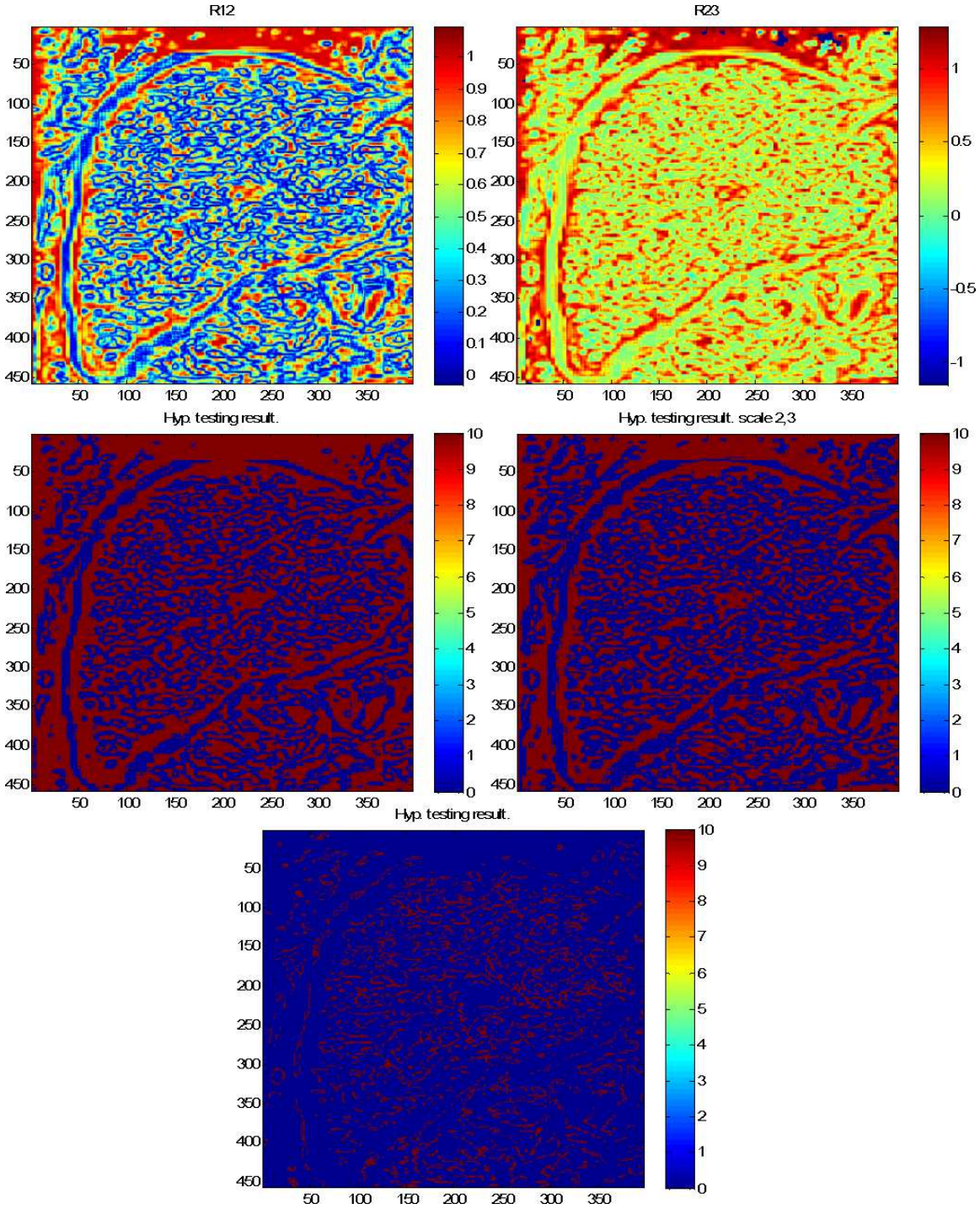


Figure 4.27: Grade 3 liver images: Fisher Z transform and hypothesis testing - first row shows the Fisher Z cross-scale correlations for scales 1-2 and 2-3; row two shows the corresponding hypothesis testing, final row is the hypothesis test result from the above three scales in total.

ple pairwise correlations to derive the structures of interest. For example, in the case of just three scales, we first calculate the cross-scale correlations and hypothesis tests for the 1-2 and 2-3 scales; then we perform a third cross correlation and hypothesis test on these two maps to combine all three scales.

The Z statistics are calculated as (Johnson, 2005):

$$Z = \frac{\sqrt{n-3}}{2} \ln \frac{1+r}{1-r} \tag{4.8}$$

where n is the sample number and r is the cross-scale correlation calculated in local neighbourhood between consecutive scales. The hypothesis test is performed at a significance level of $\alpha = 0.05$.

4.4 Conclusion

The measurements that were presented provide evidence that the Riesz weighted PC maps show more of the structural information than do the energy weighted ones. The reason behind this fact is that the local energy is inherently not suited for the representation of repeating structures. The most straightforward to understand this is in 1D, where the energy is the sum of the square of a *sine* and *cosine* function, that is it is unity. This becomes particularly evident when fine details of the presented images are magnified to show structures.

We additionally note that using a statistical approach, namely the Fisher Z cross-scale correlation with hypothesis testing, leads to meaningful quantitative PC based on monogenic LP values. Although we do not use this measure in this thesis it has several advantages over the Riesz weighted PC. The structures are much better defined and localized and also the inclusion of additional scales increases the definition of the structures instead of introducing shadowing effects as it does for the Riesz weighted PC. On the other hand, one drawback of the Fisher Z cross-scale correlation is its

computational inefficiency as compared to the Riesz weighted PC. To investigate its applicability to liver texture quantification remains as future work.

Chapter 5

Continuous Intrinsic Dimensionality

5.1 Intrinsic dimensionality

As we have noted repeatedly in this thesis, though low level image features play a fundamental role in most subsequent image processing applications, their definition is limited to the underlying geometric structure, that is to say, their intrinsic dimensionality (iD). Following this argument, it is meaningful to assign orientation to linear structures (edges and ridges), and position to 2-D localized structures, such as crossing lines or points (localized features); while homogeneous patches have neither an associated orientation nor position. Intuitively, the information content increases with the increase in the iD, and their frequency of appearance decreases in natural images, one estimate being 86% for ci0D, 11% for ci1D, and 3% for ci2D (Kalkan et al., 2005)[Sec.3., pp. 7]. This means that we expect to find far fewer significant i2D events, which is quite likely the reason why they are largely responsible for Julesz's 'pop-out' effect or saccadic eye movements in general (Krieger and Zetzsche, 1996) [Sec.II.C., pp. 1029].

iD is a non-parametric[1] description of local signal structure which differentiates it from other interesting point/region descriptors. All feature detectors introduced in the earlier chapters perform some form of iD description.

In the past, iD was defined as a hard threshold which in 2D contains three distinct labels: i0D, i1D and i2D; where i0D represents homogeneous patches, i1D are structures concentrated along a 1-D line and i2D are complex 2-D structures composed of the intersection and/or bending of 1-D structures. These structures may be also described using the local FT spectra; within this coordinate system we have an inverse relationship w.r.t. the localization property in the spatial domain: the spectra of the i0D class becomes localized to a delta peak, with the meaning of perfect predictability of the scene or localized patch. The i1D has its spectra concentrated along a line crossing the origin with a large gradient variance and small (ideally 0) line variance; finally, i2D classes have a broad frequency spread with the meaning of low predictability. The challenge is that these local structures cannot in general be classified reliably into one of the three ideal classes, because (1) measurements are intrinsically affected by noise, which imposes a nonlinear geometric deformation between the ideal signal classes (this effect is not known in advance), and (2) real signals are a mixture between these three classes in 2D with the inclusion of the possible noise component (Krüger and Felsberg, 2003), which will be the main topic of this chapter.

One particular definition of iD, for example that adopted from (Fukunaga, 1990), has a fundamental limitation in the description of local 2-D image patches. According to this "a data set in d dimensions is said to have an intrinsic dimensionality equal to d' if the data lies within a d' - dimensional subspace" (Bishop, 1995) [Sec.8.6, pp.314], (Fukunaga, 1990) [Sec.6.2., pg.280] which can be measured by their local distribution. With this formulation, curvilinear structures belong to the i1D group

[1] ciD attempts to not introduce any bias in the computations by using exclusively statistical tools in the form of variances to define the structural changes.

(Bishop, 1995) [Fig.8.11, pp. 315], which is a valid general assumption for unknown data. However, it is incomplete with respect to image features, which includes an interpretation perspective. This is because curvilinear structures are more complex than edges or lines. This is an especially important factor in the determination of image cues for the detection of tissue anomalies from the perspective of radiologists. A second, more technical, consideration is the precision of the characterization as mentioned in the introductory paragraph. Orientation may be attached to a perfect linear structure; however its confidence decreases as the respective structure deviates from linearity. Following this argument, curvilinear structures represent a transition between i1D and i2D, at least in a Cartesian 2-D coordinate system. As compared to their perfect 1-D localized counterpart, their gain of information is due to the rate of change in their tangential orientation. For images of arbitrary dimension, the meaning of an iD also has a meaning or interpretation attached and therefore the definition of features and iD poses a circular problem even for these ideal signals.

A second challenge with the definition of iDness arises from consideration of textured patches, as a separate class. In general, textures have an increased geometric uncertainty compared to the simple superposition of ideal classes and therefore the differentiation between texture and noise is, we contend, to some extent ill-posed in 2-D. This is because it is in general hard to differentiate between i0D, some i2D features, and noise. One possible interpretation of these textured classes may be as 3-D scenes; this is the fundamental position adopted in (Kalkan et al., 2007) and includes both illumination and reflectivity changes. However, such 3-D information is to a large extent lost in a single view image. A good illustration of the ill-posed nature is shown in Fig.4.(e) (Kalkan et al., 2007) in [Sec.3., pp.8]. In the context of medical images, our focus is on liver soft tissue characterisation using Magnetic Resonance Images (MRI) with the aim of detecting subtle structural changes as anatomical tissue malformations that correspond to the development of fibrosis. Compared to the anal-

ysis of natural images presented in (Kalkan et al., 2007), this would mean an inverse problem to represent what is available in the 2D images in prospect of a higher dimensional originating signal and imposes an additional constraint on the approaches to iD detection. This is because the actual decision is made based on the 2D image patch as well as depth information (Kalkan et al., 2007).

Finally, an important consideration with respect to iD detection is the reference coordinate system and the notion of feature scale. In general, iD, both without and with higher level interpretation, measures some form of variance. As a result, using conventionally available 2-D coordinate systems such measurements are bound to the space they are represented in. All features are expected to have a natural scale in which they are best described. Changing the scale of a feature inevitably changes its iDness. Using an isotropic Gaussian scale-space, with the associated general diffusion process, both 1-D and localized 2-D features locally converge to a homogeneous, i0D patch. Another example might be the change in the symmetry detection, while on the coarse scale a line may be described as a symmetric feature, with the increase of scale the two sides turn into asymmetric edges. Without listing all possible cases, we argue that in fact the definition and the interpretation of iDness of a signal depends on its local neighbourhood, and therefore cannot be accurately defined without a reference scale and coordinate system in which it resides.

As a preamble to iD estimation we introduce the formal definition of iD classes as they were presented in (Krieger and Zetzsche, 1996)[pp.1028]. In the spatial domain, the three local signal classes were defined as (Krieger and Zetzsche, 1996):

$$I(x,y) = \begin{cases} i0D & \text{if } u(x,y) \text{ is constant} \\ i1D & \text{if } u(x,y) = u(a_{11}x + a_{12}y) \text{ with } (a_{11} \neq 0) \vee (a_{12} \neq 0) \\ i2D & \text{otherwise} \end{cases} \tag{5.1}$$

The second row represents the equation of a line and u can be intensity or any

feature energy measure. In the Fourier domain this description becomes (Krieger and Zetzsche, 1996):

$$I(x,y) = \begin{cases} i0D & \text{if } U(u_x,u_y) = \text{const} \cdot \delta(u_x,u_y) \\ i1D & \text{if } U(u_x,u_y) = \frac{1}{|det(A)|} U\left(\frac{a_{22}u_x - a_{21}u_y}{det(A)}\right) \cdot \delta\left(\frac{a_{11}u_y - a_{21}u_x}{det(A)}\right) \\ & \text{with } (a_{11} \neq 0) \vee (a_{12} \neq 0) \\ i2D & \text{otherwise} \end{cases} \tag{5.2}$$

where $U(u_x,u_y) = \mathcal{FT}[u(x,y)]$, u_x, u_y are frequencies and with $\mathcal{FT}$ denoting the Fourier operator, and

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{21} \end{bmatrix} \tag{5.3}$$

is an affine transform matrix acting on (x,y) coordinates. The i0D signals that have perfect predictability, in the $\mathcal{FT}$ domain become a Dirac pulse multiplied by some constant that is proportional to the energy of the signal. The i1D features are more interesting, since they constitute a reference to the two other ideal feature classes. The definition presented in eq. (5.2) may be derived by consideration of the affine transformation of a 1D line, $u(\mathbf{Ax})$ with $\mathbf{A}$ given as the affine coordinate transform above.

Unfortunately, in both domains classification of the i2D events is unspecific (not i0D or i1D), and, as a result, several examples that fall between the iD classes end up being grouped differently by different approaches, e.g. curvilinear structures. As a result, for example, the i2D events that result from the crossing of two lines are by definition classified the same way as localized blob features, although they carry different variance information.

In general, there are several approaches to iD detection: one can investigate iDness either from the perspective of signal representation problem using a mathematical

formalism, for example, differential geometry (Zetzsche and Barth, 1990); or, from the point of view of spectral analysis (Krieger and Zetzsche, 1996); or as a combination of the two. Finally, an important consideration is whether the methodology is a detection or an estimation procedure (Krüger and Felsberg, 2003), (Felsberg and Krüger, 2003). Given that low-level features are often the first step in a higher level image analysis task and given that this step is highly under-determined, estimation of the likelihood or probability of the presence of a feature is more desired than deterministic approaches for further processing.

Using the mathematical formalism of differential geometry, the image may be modelled as a *Monge patch* $(x, y, I(x, y))$ (Krieger and Zetzsche, 1996) [pp.1030], where (x, y) are the spatial coordinates and where $I(x, y)$ denotes the image intensity. This surface may be approximated with local osculating paraboloids (Zetzsche and Barth, 1990) [pp. 1112-5], where the local i1D structure may be achieved by the bending of a 1D flat surface, which in the limit of the decrease in bending degrades to a plane, that is, the i0D class. i2D signals are suggested to be described by elliptic or hyperbolic paraboloids, since they require a local elastic deformation as compared to their 1D counterpart; this is why the design of i2D detectors becomes the most challenging task and we limit the review of the deterministic approaches to this class. Although, as we pointed out earlier, there is no agreement in the literature regarding the classification of i2D signals, in what follows curvilinear structures are classified as i2D features, characterized by their isophote curvature.

Differential i2D operators have two largely adopted one of two strategies: (1) exclusion of the i0D and i1D classes; and/or (2) the direct estimation of i2Dness through the concept of curvature. The exclusion of i0D and i1D classes requires that i2D performs a nonlinear operation and, as a result, the use of linear filters as detectors is ill-posed. This point has been stressed by (Zetzsche and Barth, 1990), who argues that this is a direct result of the *OR* operation performed by linear filters

(convolutions).

Linear filtering performs a weighted superposition (*OR* gating) of complex exponentials with a preferred orientation encoding, a formal representation of i1D signals, in which the input has been decomposed. Apart from verbal reasoning, the above limitation given the notation in eq. (5.2) may be derived as follows: an i1D signal means that at least one of u_x or u_y is not null, and for an i0D, both u_x and u_y are equal to 0. As a result, denoting the filter output as $H(u_x, u_y) \cdot \exp[j(u_x x + u_y y)]$, the only possible solution to satisfy both conditions is the trivial transfer function $H = 0$ (Krieger and Zetzsche, 1996) [eq. 16, pp. 1031].

The limitations of linear filters as i2D detectors are overcome using a nonlinear AND combination of them, while an alternative approach is the design of nonlinear polynomial filters that selectively respond to i2D features.

The first systematic approach was put forward in (Zetzsche and Barth, 1990) with two necessary and separately non sufficient conditions: (1) some form of AND operation on linear filter outputs is required, and (2) the crosstalk between the two directional derivatives needs to be minimized by the introduction of the *compensation principle*. Both of these are possible within a proper mathematical formalism using the Gaussian curvature; this is entirely determined by the determinant of the Hessian $det(H(x_i, y_j)) = I_{xx} I_{yy} - I_{xy}^2$, where I_{uu} denotes the second order derivative with respect to the dimension in the direction of u. The compensation term takes the form of the squared mixed derivatives and accounts for overlapping responses from the AND combined linear filter outputs that would otherwise lead to false responses, which is the main design criteria of i2D selective nonlinear filters.

As has been pointed out earlier, there is no agreement in the literature regarding the classification of i2D signals. Quadratic Volterra operators achieve their i2D selectivity by the non-linear combination of frequencies (Krieger and Zetzsche, 1996). As compared to linear filtering, the Volterra operator is defined with a multidimensional

convolution (Krieger and Zetzsche, 1996) [eq. 17, pp. 1031]:

$$\begin{aligned} u_2(x,y) &= h_0 + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h_1(x_1,y_1)\cdot u_1(x-x_1,y-y_1)\mathrm{dx}_1\mathrm{dy}_1 \\ &\quad + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h_2(x_1,y_1,x_2,y_2)\cdot u_1(x-x_1,y-y_1) \\ &\quad \cdot u_1(x-x_2,y-y_2)\mathrm{dx}_1\mathrm{dy}_1\mathrm{dx}_2\mathrm{dy}_2 + ... \end{aligned} \tag{5.4}$$

where $u_1(x,y)$ and $u_2(x,y)$ are the filter inputs and outputs respectively, and h_0, $h_1(x_1,y_1)$, $h_2(x_1,x_2,y_1,y_2)$, ... denote the kernels that determine the response of the system. This non-linear convolution points out the main differences as compared to linear filtering operators. The Volterra operator acts on the product (AND) of complex exponentials as compared to the weighted summation of them with linear filters. This operator includes linear convolution as a special case if the only non-zero element is $h_1(x,y)$.

The sufficient and necessary condition for the design of i2D selective Volterra operators is that all frequency pairs should be blocked that have a joint combination of the same orientation (Krieger and Zetzsche, 1996) [eq. 33, pp. 1033]; this condition is consistent with the derivations based on the Gaussian curvature and it validates both approaches. The apparently complex mathematical formulation of the Volterra operators lead to a very intuitive frequency representation; visually it accounts for the inhibition of all possible linear orientations, defined as pass- and stopbands, Fig. 5.1.

The contribution of Volterra operators to iD selection is threefold. First, by definition, they capture higher order correlations in the input signal, which has great potential for statistical signal processing[2]. Second, a more immediate application is the selective detection of different curvatures and the simultaneous detection of orientation, both of which are fundamental components of the ciD computation. The

[2]They may be considered as the nonlinear counterpart of the Laplacian (Krieger and Zetzsche, 1996) [pp. 1040]

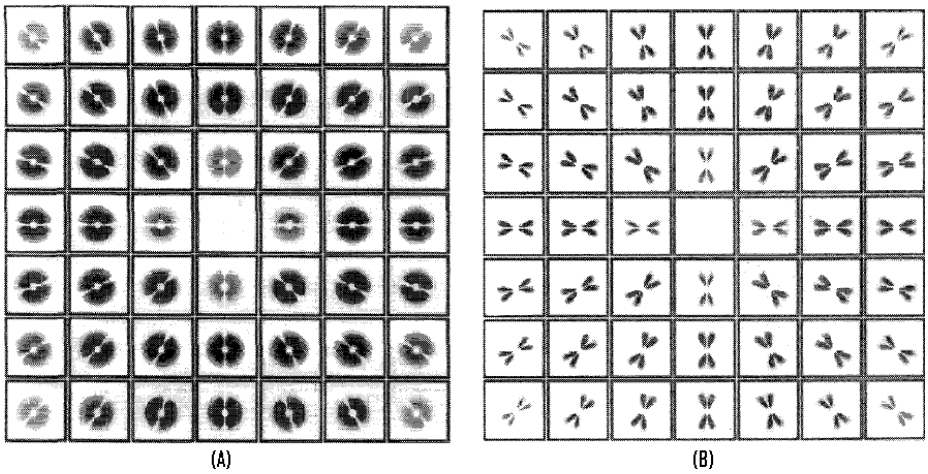


Figure 5.1: Illustration of the Volterra filters: (A) shows a Volterra filter with reduced stopband which allows the detection of small curvatures, fast changes; (B) the passbands are reduced and as a result they prefer only slowly changing curvatures. This Figure was reproduced with permission from (Krieger and Zetzsche, 1996).

curvature selective potential may be achieved by the selective tuning of the distance of the passband to the stopband, and the width of the passband. This means, that a narrow passband located in the proximity of a stopband selects 1D signals whose orientations are very close. Therefore, they allow the detection of slow orientation changes and shallow curvatures; in contrast a wide passband, that has a large distance to the stopband (to avoid the detection of slow changes) would respond to fast orientation changes, and high curvatures. For an illustration of these cases see Figure 5.1. Third, this final consideration on the curvature and orientation is exactly the frequency domain counterpart of one of the most challenging parts of the ciD computation with the consideration of all iD groups.

Apart from low level signal analysis algorithms, methodologies based on statistical learning approaches and linear approximations to the surfaces have been also considered. A difference exists between the so called global approaches in which the input data is projected into a lower dimensional space and the so called local approaches. The drawback of the global methods is that they do not preserve local connectivity information which is so important for texture measures. In the latter case, the

processing relies entirely on the local neighbourhood of the area under investigation. One interesting suggestion regarding the iD estimation was suggested in (Bruske and Sommer, 1998) in which optimally topology preserving maps (OTPM) have been used with local principal component analysis (PCA).

In OTPMs the manifold is represented by a topology preserving graph in which each node is a pointer associated with a pattern, where the neighbouring sites in the input manifold are neighbours in the graph. The essential property of OTPMs for iD estimation includes the fact that they are invariant to rotation and translation; furthermore, any embedding into higher dimensional space does not alter their performance regarding iD estimation. Local PCA approaches based on the data itself have major shortcomings for local signal characterization, also termed by the authors as the "curvature/noise" dilemma. The curvature noise dilemma is a result of local variances present on the surface. In order that the local PCA return reasonable estimates for the local eigenvalues the main criteria is that the largest variance normal to the surface must be much smaller than the smallest variance in the direction of the surface. This means that the largest variance due to noise must be smaller than the minimum intrasurface variance due to the original pattern density.

To overcome the above limitations that are caused by the direct use of the local data for the local PCA, the relative difference vectors between the central and the neighbouring elements are instead used. This estimate in fact then acts as noise cancelling mechanism over the conventional intensity based approaches, leading to a better performance. The methodology presented has two relatively weak points as compared to the ones that are being presented later in this section: the threshold based selection of the significant eigenvalues and the manual determination of the scale factor. Although this methodology is not directly applicable to our approach, it has one distinctive idea, which is to use the difference vectors from the OTPM itself for the local PCA analysis instead of operating on the intensity data patterns directly

this way obtaining more robust local PCA decompositions, which is important for the description of scenes based on local image patches.

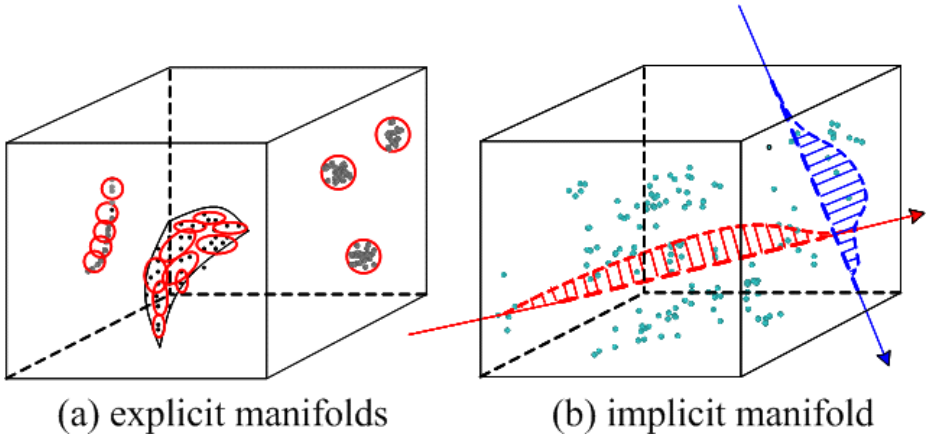


Figure 5.2: Explicit and implicit manifolds in the space of local image patches, represented by a single point. The LHS figure represent the manifolds of the ideal features, while the RHS shows the distribution of texture patches with their marginal histograms. This Figure was reproduced without permission from (Shi and Zhu, 2007).

More interesting for our purposes is the method proposed in (Shi and Zhu, 2007). Motivated by accurate object recognition based on extracted local patches the statistical approach suggested by the authors examined the identification of the mixed iD features and statistical representation of scenes that include texture and clutter (Shi and Zhu, 2007). One of the main observations made is that image patches may be mapped onto two types of subspaces that correspond to explicit and implicit manifolds. It is important to separate between these two representations because the metric by which we compare their elements may not be interchanged. The fundamental transition between the two representation may be made through a scale-space approach with the use of entropy.

The modelling framework was motivated by the observation that the so called simple signals, which we called iDs in the discussions above, have very different image generating functions to the stochastic texture patches. The first class of features (iDs) correspond to geometrically salient points or regions (Fig. 5.2, (a)), which are

highly predictable. They may be generated by functions with a small number of hidden variables (sparse coding models of geometric primitives) that describe image properties as orientation, gradient energy and photometric properties; and which may also be explicitly reconstructed. An explicit manifold may range from anything from a 0-dimensional point to 2D surfaces and above and live in a low-entropy signal world.

In turn, textured patches are suggested to be generated by Markov Random Fields (MRFs) and be represented by implicit manifolds, that correspond to a high entropy signal world. This latter approach is the one taken by most of the filter based texture description approaches to-date. The methodology proceeds by first extracting a set of features through the application of several filters and assembling the local statistics by using probability distribution functions (pdfs) and marginals (Fig. 5.2 (b)). It has been proposed that this latter method does not allow the reconstruction of the original generating function, the image patch itself may be reconstructed to a good approximation by using multi-scale descriptors, as was suggested in (Leung and Malik, 2001).

Although, it may have been overlooked in the originally suggested methodology; the observations made are highly significant for local signal characterization. The transition between the implicit and explicit manifold may be established using the notion of scale and the measure of entropy. The contribution of each manifold component to a respective sample image patch is quantified using information theoretic measures, the minimum coding length, which allows the definition of an automatic selection criteria. The simultaneous consideration of the contribution of both methodologies is measured by information gain at each respective scale.

The simulations presented in (Shi and Zhu, 2007) suggest that given a natural/arbitrary image patch the most informative scale from an information theoretic perspective is positioned between the ideal implicit and explicit manifold, where the entropy peaks. It is interesting to note that some of these observations have been

recognized earlier by the image processing community, and one of the most recent developments we would like to draw attention to being the scale-saliency algorithm presented in (Kadir and Brady, 2001).

The scale-saliency algorithm has a significant advantage compared to earlier methodologies because it achieves a simultaneous feature-space and scale-space localisation based on the measurement of unpredictability. This is important because, as we noted earlier, the notion of feature cannot be interpreted without the notion of scale. Moreover, as we also noted in an earlier Chapter, the localisation of features in both spatial and scale coordinates is represented by the so called scale-saliency map, which is defined as the product of entropy and the inter-scale saliency map $\mathcal{Y}_{\mathcal{D}}(s_p, x) = \mathcal{H}_{\mathcal{D}}(s_p, x) \times \mathcal{W}_{\mathcal{D}}$, where $\mathcal{H}_{\mathcal{D}}(s_p, x)$ denotes the entropy measure at scale s_p, spatial location x, and the quantized intensities take values from D, with $W_d(s_p, x)$ being the saliency measured over scale as $W_D(s, x) = \int_{d \in D} |\frac{\partial}{\partial s} p(d, s, x)| \mathrm{d}d$. The effective scale is selected by the peak of the entropy as $s_p = \left\{ s : \frac{\partial^2 \mathcal{H}_{\mathcal{D}}(s,x)}{\partial s^2} < 0 \right\}$. As a result, the image model described by the scale-saliency maps is restricted to the analysis of structured descriptors, and includes the elements of the iD class in some form.

However, the scale-saliency image model may be complemented with the image model suggested in (Shi and Zhu, 2007) by the inclusion of the non-deterministic component described by the implicit manifolds. This would enable one to describe the transition states between the structured feature components over scale using an information theoretic approach. We have not yet implemented this possibility.

A significant advance in iD *detection* and *estimation* was the definition of the continuous intrinsic dimensionality (ciD) topology (Krüger and Felsberg, 2003), (Felsberg and Krüger, 2003). There are two main ideas: the theoretical model with the definition of three confidence measures of the iD triangle; and a computational model that enables its description. The main hypothesis in all prior iD detection is that

there exist three different iD classes in 2D, seeking a clear separation between them. However, this assumption no longer holds for real measurements due to both geometric and noise uncertainties, and, as a consequence, a continuous transition is a better representation. There were earlier attempts at a continuous description between these classes, for example, the coherence measure for 1D-ness, with the ratio of difference and sum of the two structure tensor eigenvalues (Jähne, 1993), or the isotropy factor defined as the product of the two structure tensor eigenvalues in (Felsberg, 2002)[sec. 5.2.3, pp. 112]. Given the pairwise dependence between the three ideal classes, 1D space does not have enough degrees of freedom to represent these, and ciD is the first model that accommodates all dependencies (see Fig. 5.3) inspired by the variances of the local FT spectra. These variances then parameterize the ciD triangle (Felsberg et al., 2009).

The axis of the ciD triangle is spanned by the energy and orientation variances. These two variances may be formally measured both in the cartesian and the polar coordinate system: (σ_O^2, σ_L^2) denotes the origin, and line variance, while (σ_R^2, σ_A^2) is the radial and angular variance of the 2D local FT spectra. The origin and the radial variance measures the local grey level variation, while σ_L^2 and σ_A^2 quantifies the variance across a line, or angular variance.

These elementary measurements are correlated, but in a strict mathematical sense they are different; and are defined as (Krüger and Felsberg, 2003) [Sec.2.2, eq. (1)-(5)]:

$$\sigma_O^2 = \frac{1}{N} \int \int_\Omega |u|^2 |F(u)|^2 du \tag{5.5}$$

$$\sigma_L^2 = \min_n \frac{1}{N} \int \int_\Omega |n^T \cdot u|^2 |F(u)|^2 du \tag{5.6}$$

$$\sigma_R^2 = \frac{1}{N'} \int_0^Q q^2 \int_0^{2\pi} |F\left(q \cdot \cos(\theta), q \cdot \sin(\theta)\ \right)|^2 d\theta dq \tag{5.7}$$

$$\sigma_A^2 = \min_{\theta_0} \frac{1}{N'} \int_{\theta-\theta_0}^{\theta+\theta_0} (\theta - \theta_0)^2 \int_0^Q |F\left(q \cdot \cos(\theta - \theta_0), q \cdot \sin(\theta - \theta_0)\right)|^2 d\theta dq \tag{5.8}$$

where u is the frequency vector, Ω is the region of integration, n vector is parallel to i1D signals, Q is the radius of Ω, N' is a normalization factor, θ_0 is local orientation, $N = \int \int |F(u)|^2 du$, θ is the angle.

By definition, using the above measures, a homogeneous patch has a value equal to 0 for all of the above variances above, and it is not defined for the angular variance. Ideal 1D signals are expected to have a large σ_O^2 and σ_R^2, and 0 or low line and angular variance, depending on the levels of noise and the scale of the respective structure. The i2D features that carry the most complex information have a large variation for all four cases.

The efficient computational description of an image w.r.t. the ciD is possible with the correspondence between the second moment tensor of the local FT spectra and the structure tensor (Felsberg and Krüger, 2003) [Sec.2]. To accommodate all 2D signal classes, the ciD triangle (Fig. 5.3) is extended into the ciD cone, which in fact lifts the structure tensor approach onto a higher scale. The first coherent framework for this methodology was suggested in (Felsberg et al., 2009) and will be described in detail in the next section.

The ciD and its suggested computation provide a theoretical framework to represent structures with a continuous transition between them. In practice however, the reliability of the computed ciD values are subject to the accuracy of the input measurements and the details of the computation made at each step. To be able to quantify this 'reliability' in the remainder of this chapter, we suggest methodologies to make an informed decision at each computational steps.

5.2 Felsberg et al. (2009)'s continuous intrinsic dimensionality

As a first step of the process, structure and noise are separated based on a mixture of exponential models. Assuming a mixture of two exponentials, the probability of noise (P_{noise}) and feature (P_{struct}) is estimated at each spatial location as follows:

$$p_{mixt}(E) = \frac{P_{noise}}{\mu_{noise}} \exp\left(-\frac{E}{\mu_{noise}}\right) + \frac{P_{struct}}{\mu_{struct}} \exp\left(-\frac{E}{\mu_{struct}}\right) \tag{5.9}$$

where E denotes feature energy and μ is the mean of the respective exponential distribution. Then, using a sigmoid function, the image energy is mapped onto the interval $[0, 1]$. Estimation of the mixing coefficients and the exponential mean values is performed using Expectation-Maximization (EM). Using these estimates, the energy map is soft-thresholded using the model of activation functions (Bishop, 1995) with the model:

$$d(E) = \frac{p(E|C_{struct})P(C_{struct})}{p(E|C_{struct})P(C_{struct}) + p(E|C_{noise})P(C_{noise})} \tag{5.10}$$

Which, given the mixture of exponentials, is a sigmoid

$$d(E) = \left(1 + \frac{P_{noise}\mu_{struct}}{P_{struct}\mu_{noise}} \exp\left(E \cdot \left(\frac{1}{\mu_{struct}} - \frac{1}{\mu_{noise}}\right)\right)\right)^{-1} \tag{5.11}$$

with the assumption that $\mu_{noise} < \mu_{struct}$. In the rest of this chapter we will use the same notation for the image energy (E) without making note of the above mapping and we proceed with the methodology itself.

Using the couple (E, θ) the double angle representation is built $h(E, \theta) = E \cdot e^{i2\theta}$. This corresponds to the mapping onto the circle parameterized by $(R, \theta) = (E, \theta)$, and in the cone model represents a cross section of the cone parallel to the

base. The uncertainties for E and θ are indicated by the $[E-\delta_E, E+\delta_E]$ and $[\theta-\delta_\theta, \theta+\delta_\theta]$ intervals, the sources of them being: (1) initial estimate of (E, θ), (2) the estimated (P_{noise}, P_{struct}) and the decision made between ideal ci1D and ci2D classes. Coordinates derived from these allow the positioning within the cone with the coordinates $(E, E\cos(2\theta), E\sin(2\theta))$.

The decision between the 1D and 2D features is based on the assumption that the weighted local average of h in the spatial domain quantifies the coherence in orientation and yields an estimate of the 1D structure versus the 2D. In practice this is performed by using a Gaussian weighted average on each of the coordinates.

To relate this to tools currently popular ones in computer vision, a parallel may be drawn with part of the histogram of oriented gradients (HOG) (Dalal and Triggs, 2005).

Examining the two methods, there are several conclusions that may be drawn. The local averaging on the cone coordinates proposed in (Felsberg et al., 2009) is essentially a summation over the HOG histogram bins weighted by their associated orientation vector. This means that all structural information about the local histogram is discarded in the suggested method (Dalal and Triggs, 2005). Second, the double angle representation that is used in (Felsberg et al., 2009) provides a proper mathematical framework to describe the HOG histograms. Third, yet another way to think of the ciD cone is in terms of linear diffusion. The scale increases from the base to the top of the cone, which agrees with the cardinality of the structures to be represented with scale changes. In this respect the ciD cone may itself provide the mathematical framework for the description of the hierarchy of the HOG histograms, with the note that their respective roles are different. We will not develop this aspect here, rather we leave it for further work. The reason we mentioned this aspect is simply to point out that a rich source of information in form of the local HOG pdfs is overlooked in the current HOG model.

Finally, a point that is highly localized will tend to be positioned in the centre of the double angle circle, while a structure with high orientation coherence will be closer to the surface of the cone. As a final step in the proposed method, the cone coordinates are collapsed onto the *ciD triangle* shown in Fig. 5.3 and the ciD probabilities identified using the barycentric coordinates as the area of the triangle opposite, each of which is bound to uncertainties.

There are several sources of uncertainty in the accurate definition of the *ciD* triplet. We distinguish between two sources:

1. *input measurement error*: of the input image energy (E) with local orientation estimate (θ), and

2. *process errors*: which are due to the methodology.

In medical imaging both of these may be reduced by making an informed choice tailored to the application in terms of feature detection.

5.3 Proposed methodology

This section describes the approaches we have taken to make more informed decisions on the computational aspect of the *ciD* probabilities and to derive their associated uncertainties.

5.3.1 Uncertainty description and propagation: *all measurements are associated with a pdf*

Mathematics provides an exact description of structures and shapes; these idealized representations are brought into the real world of applications by the notion of *uncertainty*. This means that we accept that all measurements are prone to errors; the

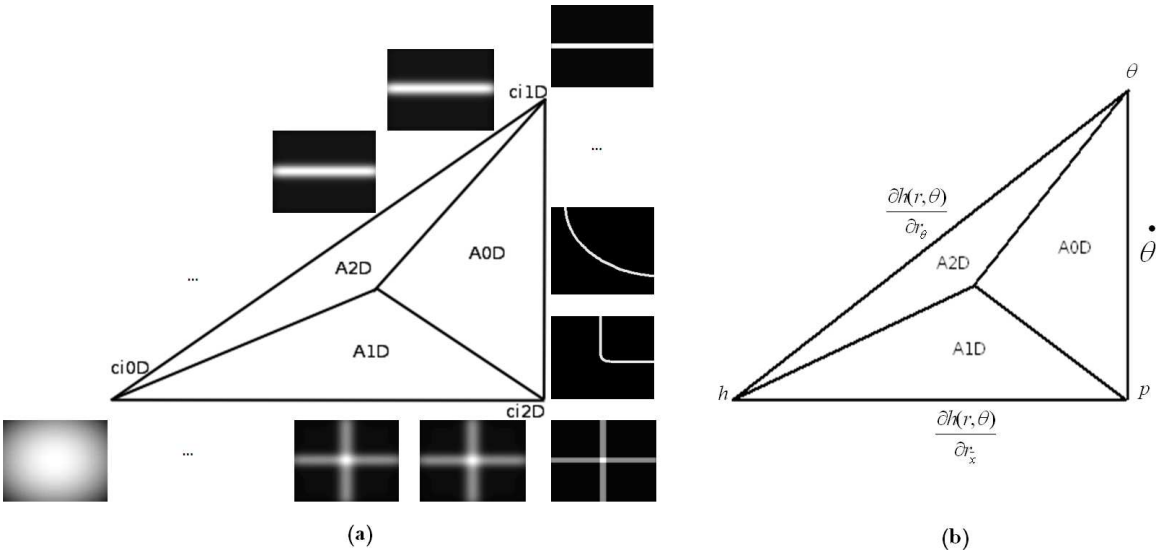


Figure 5.3: Continuous intrinsic dimensionality triangle. The figure (a) was reproduced after the illustrations shown in (Felsberg et al., 2009). The vertices denote the ideal signal classes: homogeneous, i2D an i1D. The edges of the triangle allow the transition between these ideal classes as follows: ci0D-ci1D continuous increase in the line profile, ci2D-ci0D smoothing of an ideal 2D signal - decrease in energy, ci1D-ci2D increase in curvature - decrease in orientation coherence. Figure (b) contains the physics relations that allow to define the uncertainties.

engineering approach is not only to try to reduce these errors, but also to predict them.

Our aim is to estimate the uncertainties associated with the derived ciD probabilities based on the estimated error in measurements in the input energy and orientation map, and knowledge of the ciD models used. The general parameterizations of the uncertainties in the ciD framework, both for the spatial and cone coordinates, are presented in Fig. 5.4. In what follows, we use the definitions introduced in (Pennec and Thirion, 1997), (Pennec and Ayache, 1998), (Pennec, 2006), while we derive the respective formulae for the propagation of uncertainties. More specifically, we adopt the formalism described in (Pennec and Thirion, 1997) [pg. 205-206].

In brief, as the title of this section states, all measurements are best described by a pdf that accounts for the uncertainties present in the system. For computational reasons, we only retain the first and second order centred moments in the form of

the expected value (or expectation matrix) and the covariances with the assumption of a Gaussian distribution. That is, at each spatial location the orientation and the energy are described by the couple $\left(\overline{\theta}, \Sigma_{\theta\theta}\right)$ and $\left(\overline{E}, \Sigma_{EE}\right)$, where the $\overline{x}$ denotes the expectation values and Σ_{xx} are the covariances defined as:

$$\overline{x} = E(\hat{x}) = \int y \cdot p(y) dy \tag{5.12}$$

$$\Sigma_{xx} = E\left((\hat{x} - \overline{x})(\hat{x} - \overline{x})^T\right) \tag{5.13}$$

The propagation of the uncertainties through an explicit function follows eq. (1) in (Pennec and Thirion, 1997) [pg. 205], in which the first order approximations by Jacobians (J) are used for the propagations of the uncertainties. Given a bivariate function $\overline{z} = h(\overline{x}, \overline{y})$, where $\overline{x}$ and $\overline{y}$ are independent, the covariance matrix is estimated as:

$$\Sigma_{zz} \quad = \quad J_{h_x} \Sigma_{xx} J_{h_x}^T + J_{h_y} \Sigma_{yy} J_{h_y}^T \tag{5.14}$$

where the notations are as described above. This definition and its generalization is applied throughout the derivation of the uncertainties for the *ciD* triplet.

5.3.2 ciD framework uncertainties

As described earlier, uncertainty within the *ciD* framework derives from two main sources: *measurement errors* associated with the orientation and energy map; and *process uncertainties* that are due to the specific choice made in each computational step.

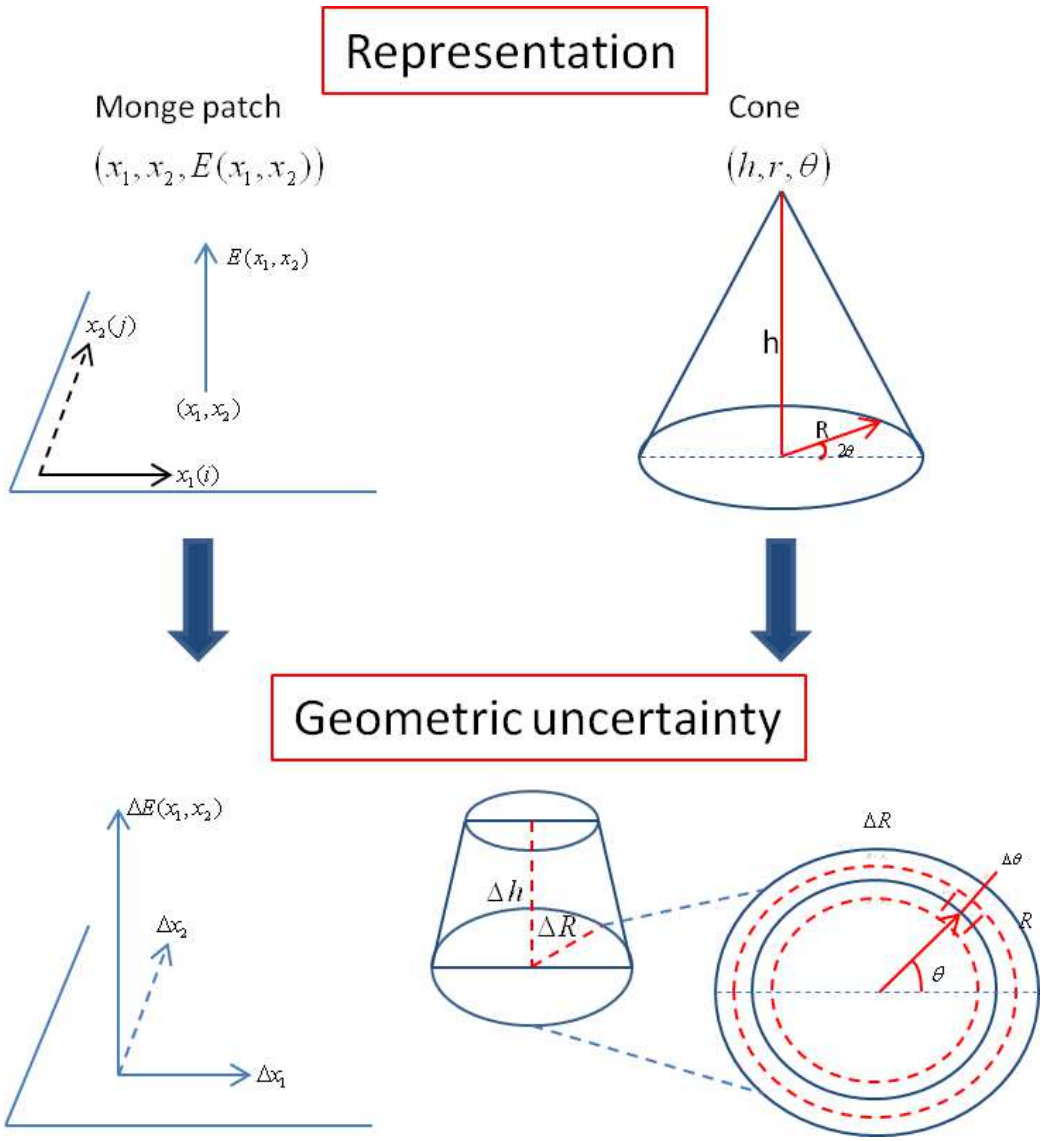


Figure 5.4: Image parametrization and uncertainty representation. On the left hand side (LHS) of the figure the conventional spatial domain Monge patch is shown, while the right hand side (RHS) shows the ciD cone representation. The geometric uncertainties that are shown below illustrate that each representation carry three sources of uncertainty. In the spatial domain we have three sources of measurement errors: displacement in the 2D plane, denoted by $(\Delta x_1, \Delta x_2)$, and measurement error in the accuracy of the image energy, or intensity, $\Delta E(x_1, x_2)$. In the cone, the most natural way to account for the uncertainties is by the use of their displacement of each of the coordinates. The truncated cone represents the geometric uncertainty within the cone coordinates.

Input measurement error

Quantification of the *measurement error* is especially important because this is the foundation of the *ciD* cone. The associated uncertainties may be reduced and mod-

elled for specific applications. However, the choices are highly interdependent and often restricted to the available mathematical formalism. More specifically, first, based on the type of structures present in the images, the estimation methodology is set; second, a statistical model is chosen to describe the local distribution of the orientation and energy. Noise throughout this description is isotropic and varies globally within the image, however, it has a local effect on the specific distributions.

Statistical models for the local characterisation of orientation and energy are inevitable. The first step is the choice between a parametric and non-parametric distribution. Parametric distributions are best used when there is strong evidence about the expected image tissue model, so the description with low sample sizes may be more accurate. However, these distributions also include a bias and are limited when describing general pdf shapes. Non-parametric models, especially those that build on the local covariances of the structure, better describe the local signal, but their use is limited. Often, in medical imaging, parametric distributions to describe the imaging modality, for example ultrasound, is available, and therefore computational methodologies focus on these.

There is no general parametric model to describe all the shapes that local orientation and energy distributions show in response to the local structures available in the images. From this point of view, we differentiate between flat and peaked, then additionally symmetric and asymmetric distributions. It is our observation (based on the estimation methodologies of choice), that none of the available distributions can accurately describe such changes. The particular distributions that we have examined include the Nakagami $\mathcal{N}(m, \omega)$, Normal $\mathcal{N}(\nu, \sigma)$ and Exponential $Exp(\lambda)$ family.

For our purposes it is desired to use a mixture of $\mathcal{N}(m, \omega)$, $Exp(\lambda)$, and $\mathcal{N}(\nu, \sigma)$. The reason behind this choice is that the Nakagami distribution is able to represent a wide variety of shapes, the Exponential distribution is the descriptor of a Gaussian noise field and the Gaussian is preferred for computational considerations. The

Nakagami distribution, defined as (Destrempes and Cloutier, 2010):

$$p(x|m,\omega) = \frac{2m^m x^{2m-1}}{\Gamma(m)\omega^m}\exp\left(-\frac{m}{\omega}x^2\right) \tag{5.15}$$

is capable of exhibiting a rich set of shapes (m) and a scale (s) variation that is to move an 'exponential like' decay to a symmetric 'Normal like' distribution. However, it is a flat distribution and the above formalism can not describe either a low variance exponential or a highly peaked Normal distribution. There are ways to alleviate this to some extent. We argue that these specific cases that we term *extreme shapes* both in their symmetric and asymmetric form often arise in the proximity of an ideal feature, therefore their presence cannot be ignored as such.

The Exponential and Normal distribution families are the most convenient sources to represent the *extreme* pdf shapes. The exponential distribution defined as (Johnson, 2005) [pg. 170]:

$$p(x|\beta) = \frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right) \tag{5.16}$$

where the β parameter controls both the mean and the variance of the distribution. The univariate Normal distribution is defined as (Johnson, 2005) [pg. 154]:

$$p(x|\nu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-(x-\nu)^2}{2\sigma^2}\right) \tag{5.17}$$

where ν represents the mean and σ the standard deviation of the distribution. Both theoretically and computationally these two distributions can be fitted to highly peaked distributions.

Finally, we might consider using non-parametric distributions and histogram based approximations. The general description of their shapes and scales are however limited to the measurement of symmetry and peakedness of the respective distribution.

Peakedness is measured by the fourth order moment of the distribution better known as *kurtosis*:

$$Kurtosis = \frac{\sum_{i=1}^{N} (x_i - \overline{x})^4}{(N-1)s^4} \tag{5.18}$$

where x_i denotes the ith sample, $\overline{x}$ is the sample mean and s the standard deviation. The formal assessment of symmetry is done using the third order moment, the *skewness*:

$$Skewness = \frac{\sum_{i=1}^{N} (x_i - \overline{x})^3}{(N-1)s^3} \tag{5.19}$$

where the notations are the same as for the kurtosis.

The mapping that is applied to move from the unbounded double angle representation to the cone coordinates requires that a decision between the structure and noise be made. Therefore, noise cannot simply be accounted for as part of the uncertainties describing the input maps.

Noise estimation in (Felsberg et al., 2009) was based on the assumption that both structure and noise can be modelled as exponential distributions. Indeed, the squared gradient magnitude of a Gaussian noise field is exponentially distributed (Förstner, 1998) [pp. 9.]. However, the distribution of squared gradient magnitude of the structures vary in a way that depends on the underlying structure (Fig 5.5). The figure shows that the Exponential distribution for the description of the structure energy holds for natural textures (green line). In turn, the Gamma distribution is more general and it includes as a special case the Exponential. In general, the statistics of natural images were examined in several studies and it has been suggested that they follow a Weibull distribution (Geusebroek, 2005). Based on our observations,

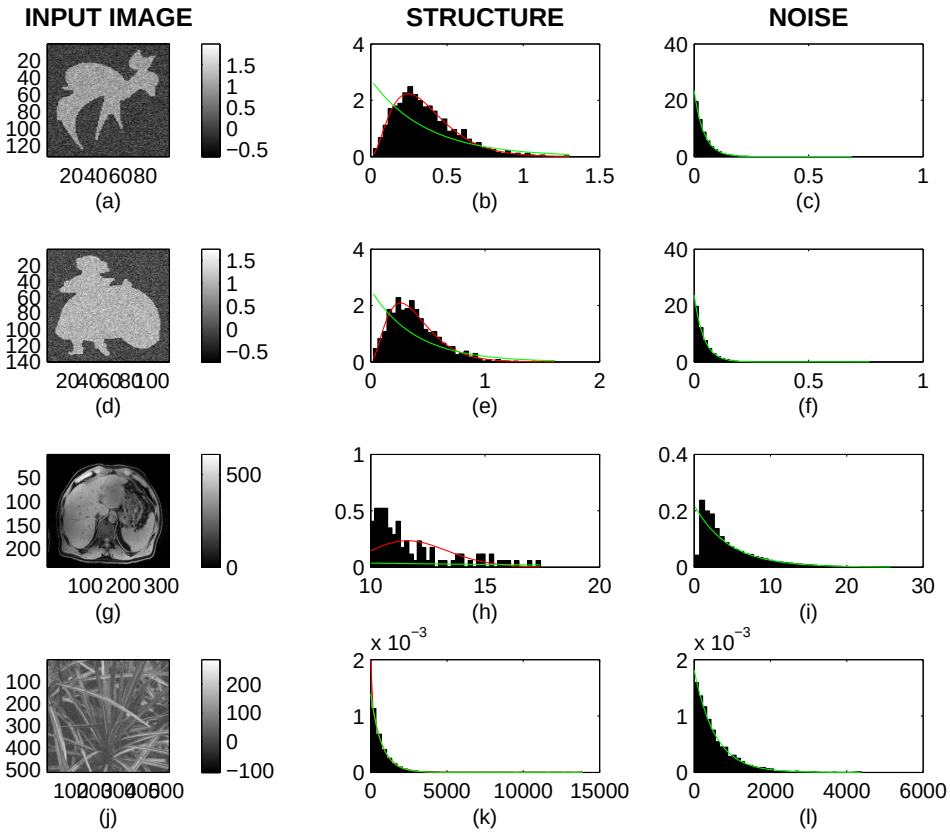


Figure 5.5: Structure and noise statistical modelling. The columns from left to right are as follows: the input images, where (a) and (d) are samples from KIMIA-database (2011) and (g) is a liver MRI and (j) is a natural texture from the IM VIS Texture database, the second column shows the distribution of the structures, Fig. (b), (e), (h), (k), and finally the last column shows the noise distribution. The curves on the top of the normalized histograms represent the estimated Exponential (in green) and Gamma distributions (red). The Gamma distribution could not be fitted to the highly concentrated exponential distributions therefore they are not present on these images.

on binary images and naturally textured images with added Gaussian distributed noise, the structures are well approximated by the Weibull, Lognormal and Gamma distributions, whereas the Exponential assumption only holds for a limited number of samples.

The assumption of the exponential mixture model for the separation of structure and noise (Felsberg et al., 2009) holds for 1D structures, while there is a systematic

change towards a Gamma like distribution with increased curvature. This observation was made using additive Gaussian noise, and therefore may not be applicable to general noise processes. We do not currently understand the reasons behind these distributional changes, though we have yet to conduct an exhaustive investigation of this as a function of the isophote curvature.

Based on these experimental observations, our hypothesis is that the mixture of exponential model is valid only for a limited set of image data and that in general the structure component is closer to a log-normal or Gamma distribution. As we pointed out earlier, the Gamma distribution has the advantage over all other distributions that it includes the exponential as a special case, therefore there is no need to compromise in the representation of the background or noise component.

The mixture of Gamma distributions is better suited to model the structure and noise on squared gradient magnitude images. In order to demonstrate this, Figure 5.6 shows quantitative measures between the empirical data set derived from the KIMIA database images (KIMIA-database, 2011) and the model predicted distributions. Figure (a-d) shows statistics for the same 10% noise level over the multiple images, while Figure (e-h) presents statistics over the same sample image with increasing additive noise effects. Figure (a-c) shows that the largest sum of absolute distances (SAD) between the PDFs and CDFs and the KS statistics is because of the misfit of the exponential distribution to the structure component (EXP-S) and the lowest distance due to the exponential model noise model (GAMMA-N, EXP-N). Note that the fact that the structure component fitted to the Gamma distribution is situated in the middle is not surprising given that the Gamma distribution does not give an exact fit to the experimental structure. The p values derived from the KS statistics confirm this. The highest p values are achieved with the fit to the noise component whereas the lowest p values are due to the effect of the misfit between the empirical structure component and the exponential model. The difference between the Gamma and ex-

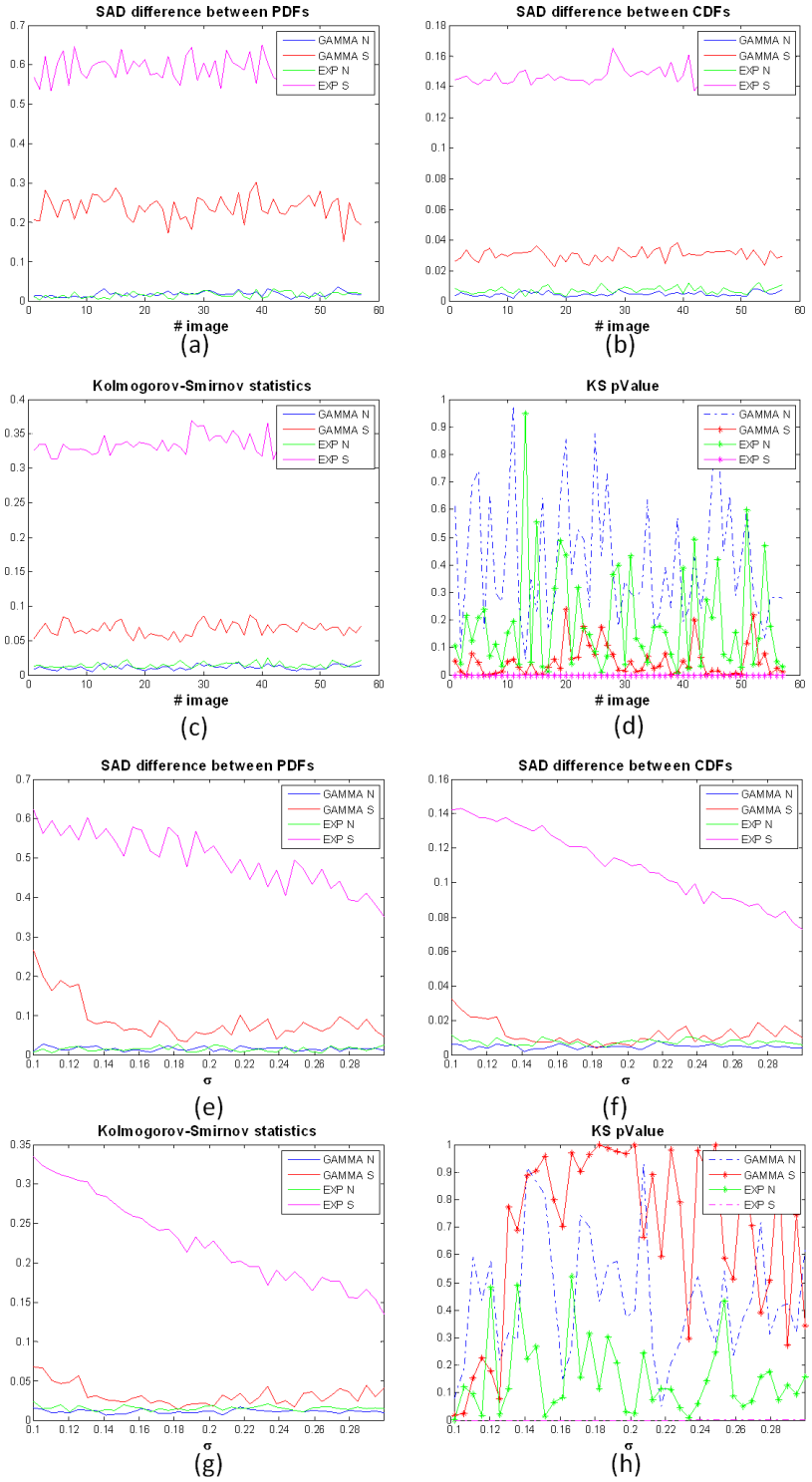


Figure 5.6: Comparison between using the mixture of Gamma or Exponential distribution for modelling the structure and noise component in the squared gradient magnitude image. Due to space limitations please see the text for the explanation.

ponential fitted component is due to the fact that the random noise model use for the Exponential and Gamma experiments do not coincide.

Similar observations may be seen in figures (e-h), however, the change over increasing noise effect shows an improvement in the accuracy of fit. This effect was also observed in a set of experiments conducted using the Hidden Markov Measure Field model. Though seemingly contradictory, this is because the squared gradient magnitude distribution of a Gaussian noise field corresponds to an exponential distribution (Förstner, 1998). As a result, with the increasing contribution from the Gaussian noise, the distribution of the structure component converges to an exponential. This has been documented earlier (Förstner, 1998)(although not for structure modelling) so we refrain from presenting the detailed mathematical derivations. What is obvious from Figure (f) is that despite the apparent accuracy with increasing additive noise as detailed earlier, the Gamma distribution achieves far better description of the structure distribution as is suggested by the p-values derived (GAMMA-S, red).

Given a parametric noise model, which is often the case in medical imaging, the challenge is to estimate the parameters of the distribution ideally by accounting for the local spatial connectivity of the pixels.

Process uncertainty

The (Felsberg et al., 2009) model has two main sources of process uncertainty:

1. *Noise parameter estimation* and choice that maps down the energy measure to the cone coordinates, and

2. *Informed decision* made between the ideal transition between a $ci1D$ and $ci2D$ class features.

The choice of the noise and structure means that the differentiation between the $ci0D$ and $ci2D$ classes derives from the extent to which they are locally spatially

correlated. Based on the requirements in the original formulation, we choose to use Hidden Markov Measure Fields (Marroquin et al., 2003), since they have the following advantages:

1. In the estimation of the global noise parameters we can elegantly take into account the spatial connectivity of voxels;

2. It allows control over the weights of the hidden p-field (class probability); and

3. Parametric noise and structure models are often available for the specific imaging modalities, which can be effectively incorporated as prior probabilities into the given iterative framework.

In deriving our methodology we focus on ideal features.

The transition between the $ci1D$ and $ci2D$ class is not only determined by the uncertainties that come from the input measurements, but also the assumption built-in to the computations. In Section 5.2 we already drew attention to the extent to which the information that is overlooked in the original estimation step with the suggestion that HOG (Dalal and Triggs, 2005) might be a better way to characterize these transitions. This is based on the assumptions that the conceptual difference between the three ideal iD classes includes:

1. $ci0D$ class includes low or no energy and a high scattering of orientation due to the presence of noise. As a result, the local energy pdf is characterized by a $\delta(t-u)$ peak in the noiseless case; and a flat distribution in the presence of noise. The orientation variance present is expected to be due to noise. We measure this as homogeneity, meaning that there is no orientation or energy localisation for this class.

2. $ci1D$ is expected to be a high energy and highly peaked spectra on the 1D HOG histograms, due to the fact of the coherence of the orientation values

are positioned along the line. The ideal 1D line can be characterized by its orientation, with the meaning of high orientation certainty.

3. $ci2D$ in turn is expected to have a peaked energy distribution and a flat orientation distribution. As a result, it can be localized in position. In relation to the 1D $ci0D$ HOG histogram we may have a similar shape compared to the $ci0D$ class, however, with a clear distinction between the associated energy distribution for the non-normalized histograms.

With these characterizations, the non-normalized HOG histogram might be a good description for the different ciD classes and the transitions between them.

The 2D extension of HOG is a more informed way to describe the shape changes within the ciD triangle. This does not imply a valid joint histogram, rather we mean that instead of the summation of the energy values, each local orientation bin is associated with a 1D energy histogram. To examine whether the above shape changes are a valid formalism, we examined whether they consistently map to the same area of the ciD triangle.

ciD uncertainty propagation

The propagation of uncertainties in the ciD framework is represented in Fig. 5.7. In practice, there are two explicit functions through which the uncertainties are propagated. The three orthogonal coordinates that define the coordinates of the cone are defined as $(d_1, d_2, d_3) = (|d|, \Re(d), \Im(d))$ after performing the Gaussian weighted local averaging of the double angle vectors in the spatial domain.

Following the process described in (Pennec and Thirion, 1997) [pg. 205] there are three main steps by which the uncertainties and mean values of the parameters are calculated. These propagation steps include the uncertainties carried through the sigmoid function, the estimation of the double angle coordinates, the convolution with a Gaussian $G(0, \sigma)$ function to make a decision between the $ci1D$ and $ci2D$ classes,

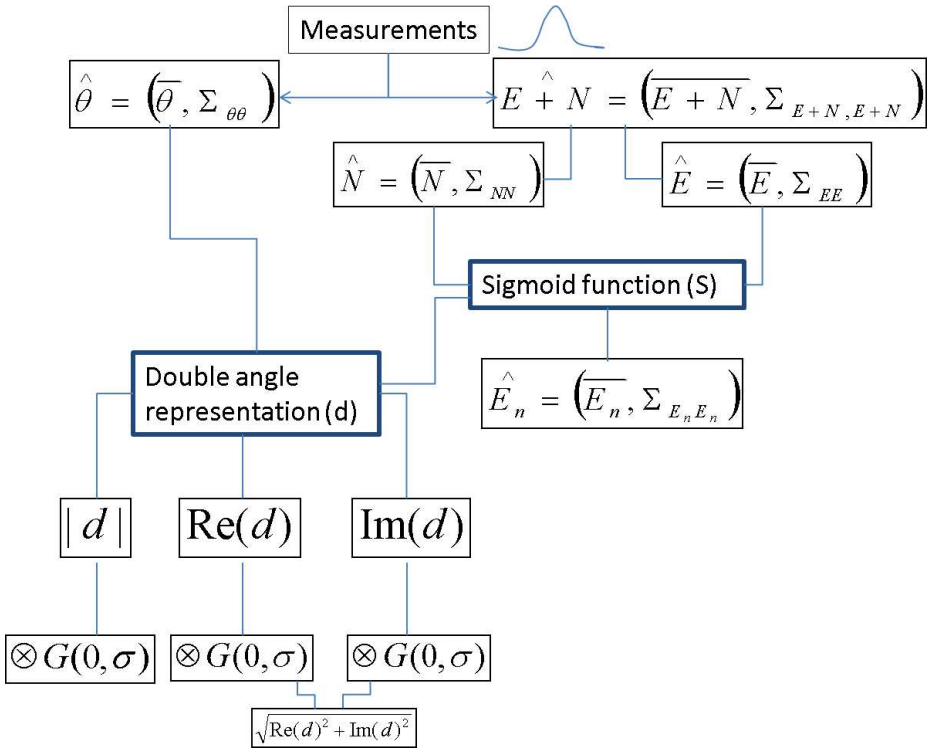


Figure 5.7: Schematic representation of the uncertainties within the continuous intrinsic dimensionality. The first step in this process is the estimation of the local orientation (θ), energy (E) and noise (N). Each measurement position is associated with a mean and a covariance. Within the *ciD* framework, the sigmoid function is applied to map the double angle coordinates onto the cone. Given the measurements associated with the orientation and local energy the position within the cone is calculated by local averaging the double angle coordinates in the spatial domain. The practical computations follow the parametrization of the cone with three orthogonal axes, corresponding to the magnitude, real and imaginary parts of the double angle representation after local averaging.

the associated functions of these are denoted by d, h and f_d, and finally the mapping onto the cone coordinates. First, we calculate the mean and the covariance of the energy as:

$$\Sigma_{E_nE_n} = J_{d_E}\Sigma_{EE}J_{d_E}^T + J_{d_N}\Sigma_{NN}J_{d_N}^T \tag{5.20}$$

$$J_{d_E} = \frac{P_2}{\lambda_2^2 P_{mixt}^2}\exp\left(-\frac{E}{\lambda_2}\right)\left[-P_{mixt} + \frac{P_2}{\lambda_2}\exp\left(-\frac{E}{\lambda_2}\right)\right] \tag{5.21}$$

$$J_{d_N} = \frac{P_1P_2}{\lambda_2\lambda_1^2}\exp\left[-\frac{E(\lambda_1+\lambda_2)}{\lambda_1\lambda_2}\right]\frac{1}{P_{mixt}^2} \tag{5.22}$$

where $\Sigma_{E_nE_n}$, J_{d_E}, J_{d_N} denote the updated energy and noise covariance matrix and the respective derivatives of the Jacobians. The model parameters follow the convention used in Sec. 5.2, where the index 1 and 2 denote the noise and the structure parameter respectively. Following the same convention, the mapping from the double angle representation results in the covariance defined as:

$$\begin{aligned}
\Sigma_{dd} &= J_{h_E}\Sigma_{EE}J_{h_E}^T + J_{d_\theta}\Sigma_{\theta\theta}J_{h_\theta}^T && (5.23)\\
J_{h_E} &= \frac{\partial h\,(E,\theta)}{\partial E} = \frac{\partial}{\partial E}\left(E\cdot\exp(i2\theta)\right) = \exp(i2\theta) && (5.24)\\
J_{h_\theta} &= \frac{\partial h\,(E,\theta)}{\partial \theta} = \frac{\partial}{\partial \theta}\left(E\cdot\exp(i2\theta)\right) = i2E\cdot\exp(i2\theta) && (5.25)
\end{aligned}$$

where the notations are as for the earlier step.

With a spatial model the propagation includes a spatial convolution on the three components with a Gaussian kernel, and we denote each of these functions as f_d defined as: $f_d = \sum_i \sum_j d(n,m)G_\sigma(n-i,m-j)$, where σ denotes the standard deviation of the smoothing kernel.

The uncertainties from the three components $(|d|, d_1, d_2)$ are propagated onto the triangle coordinates (x_t, y_t) as:

$$\begin{aligned}
(\overline{x_t}, \Sigma_{x_tx_t}) &= \left(\overline{E}, \Sigma_{EE}\right) && (5.26)\\
(\overline{y_t}, \Sigma_{y_ty_t}) &= \left(\frac{\overline{d_1}^2 + \overline{d_2}^2}{\overline{d}}, J_{y_td}\Sigma_{dd}{J^T}_{y_td} + J_{y_td_1}\Sigma_{d_1d_1}{J^T}_{y_td_1} + J_{y_td_2}\Sigma_{d_2d_2}{J^T}_{y_td_2}\right)
\end{aligned}$$

These uncertainties are subject to uncertainties with respect to geometric constraints.

5.3.3 Orientation, image energy and noise estimation

ciD is designed to incorporate all 2-D signals and the accurate representation of it by the shape of parametric distributions is currently limited. It is desired that the

specific distribution of choice is capable of representing both symmetric and asymmetric shapes with a continuous transition between them. Our main interest is in the accuracy of the predicted mean and covariance values for our computations. As a result, to find the distribution that best describes the sample orientation we reflect the distributions and use the linearity of the mean to estimate the original distribution's mean.

Local orientation estimation and statistical description

We compared three state-of-the-art orientation estimation methodologies using as measure the estimated mean with their confidence and the relative error distribution.

Both an ideal 1-D feature and curved line was used to examine our choice for orientation estimation and modelling the local distribution. The ideal 1-D feature was modelled as a Gaussian line profile $I(x,y) = \exp^{-\frac{(x\cdot\mathrm{Cos}(\frac{\pi}{2}-\theta)-y\cdot\mathrm{Cos}(\frac{\pi}{2}-\theta))^2}{2\cdot\sigma^2}}$, and the curved line as the circle approximated by Bresenham's algorithm (Pitteway, 1967).

The orientation estimates from the 1-D rotating line experiment were collected from the line profile area. The selection of local LO patches to examine their distribution needs more consideration both as regards the coordinate frame and the actual shape of the mask. The rate of change of orientation along the edge is a function of the circle radius that can no longer be neglected, $\Delta\theta \nrightarrow 0$. There are several options for the choice of the local mask shapes (1) rectangular mask, (2) *circular* mask, and ideally to use a polar coordinate framework in which the selected range of the circle covers an interval $[\theta - n\Delta\theta, \theta + n\Delta\theta]$, where θ is the centre of the selected interval, and the central part matches the local curvature of the arc. Due to computational consideration orientation values were collected from an area of circular patch.

The results of the orientation estimation step are presented in Fig. 5.8- 5.9. Three orientation estimation methods are compared: the Gradient Energy Tensor (GET) (Felsberg and Köthe, 2005), the Monogenic Signal (MONO)(Felsberg and Sommer,

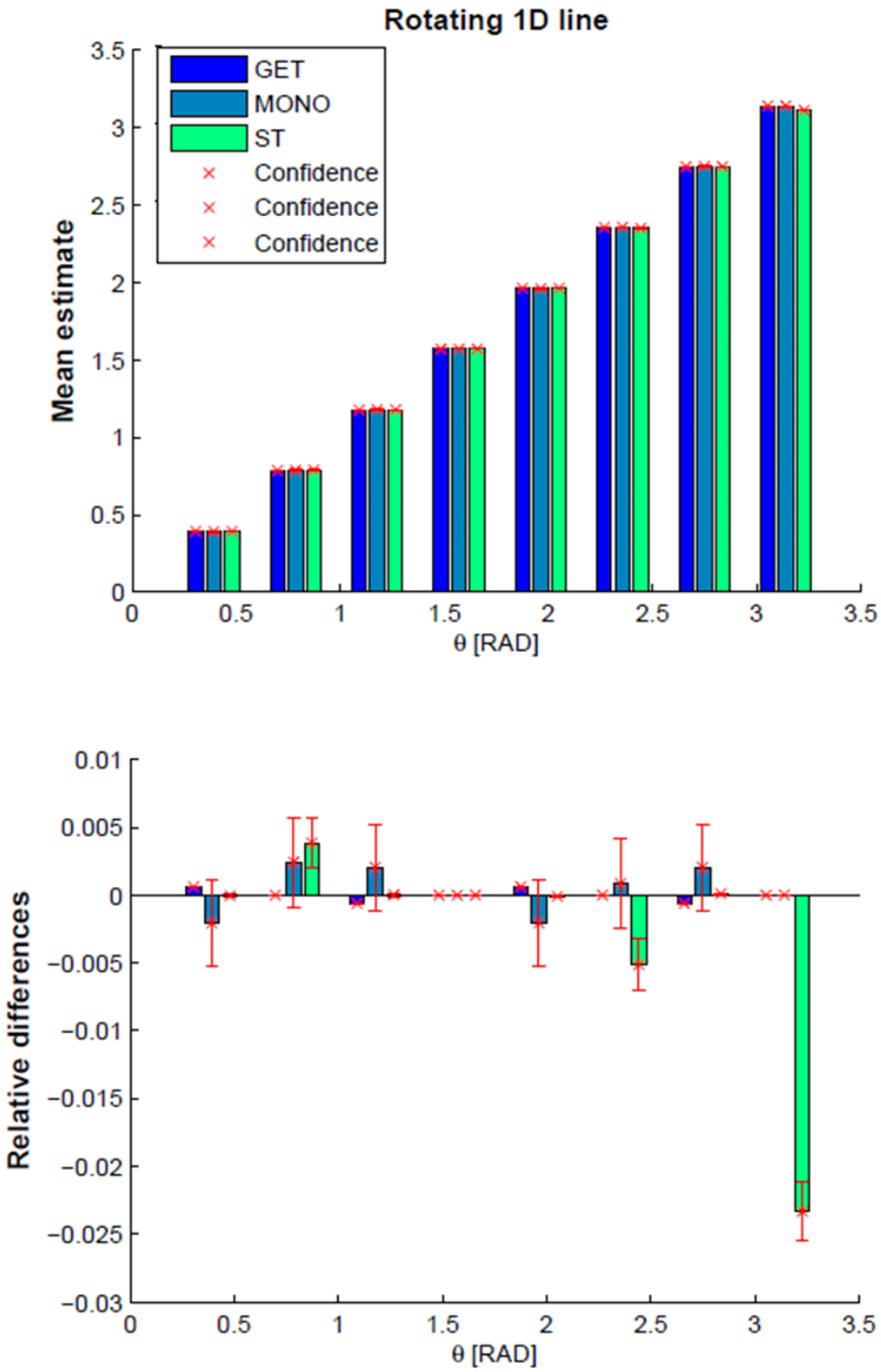


Figure 5.8: Rotating line in 2D: comparison of state-of-the-art orientation estimation methodologies: GET, Monogenic Signal and Structure Tensor. The top figure shows the mean orientation estimate, while the bottom figure the relative differences as a function of orientation angle. Both barcharts contain the confidence bars based on the assumption of a Gaussian process.

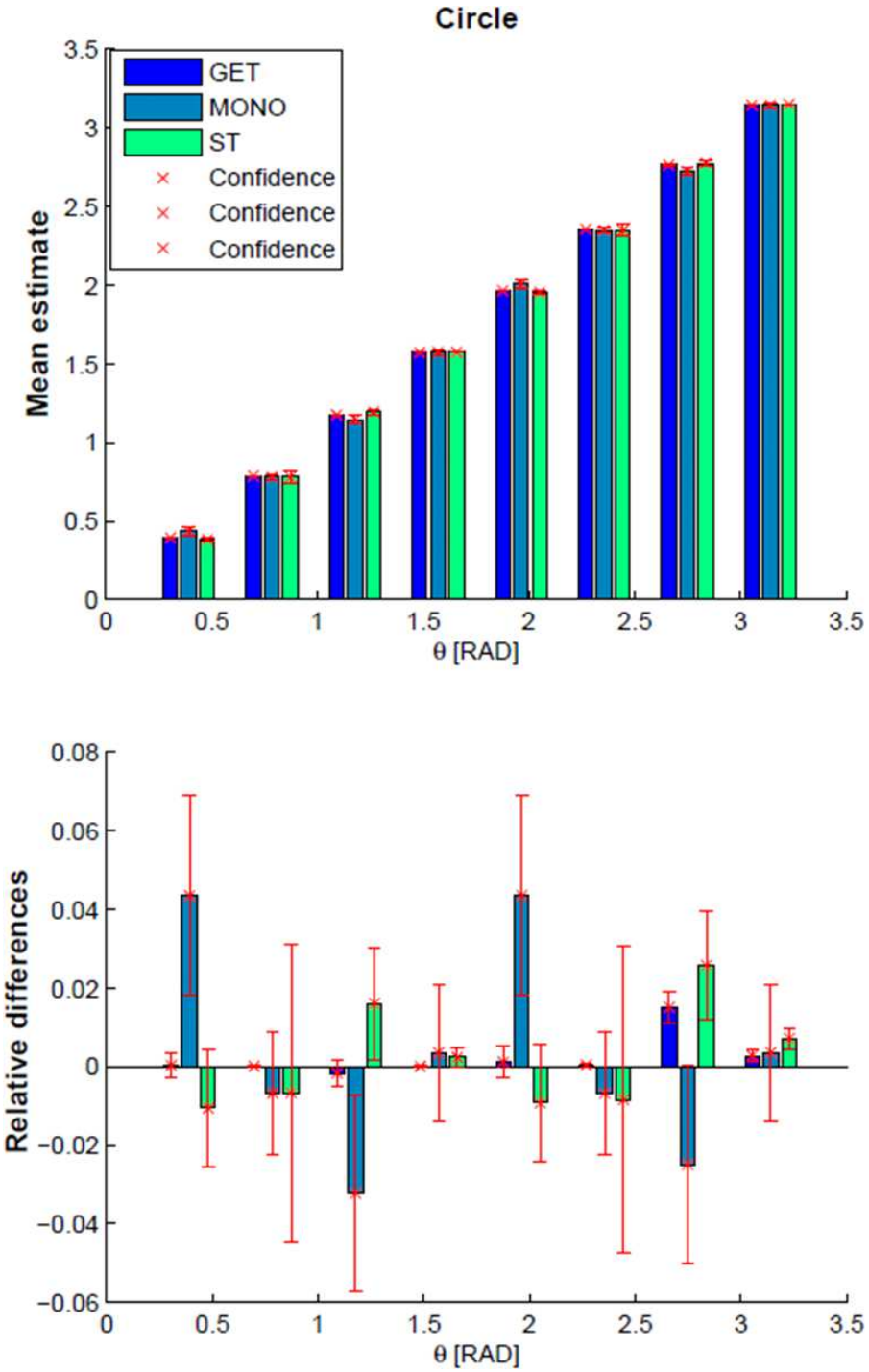


Figure 5.9: Circle: comparison of state-of-the-art orientation estimation methodologies: GET, Monogenic Signal and Structure Tensor. The first statistics show the mean orientation estimates, while the second shows the relative differences between the estimated and the ideal orientation. Both barcharts contain the confidence intervals based on the Gaussian assumption.

2002), and the structure tensor (ST) with the Scharr filter(Weickert and Scharr, 2002). These estimation methodologies have been already described in the literature review chapters, except for the GET operator which is illustrated in Figure 5.10. The top two figures show the estimated mean orientation values with their confidence bounds, while the bottom two figures show the relative difference estimated means. Ideally the expected mean value in the latter scenario is zero. The error bounds show the 95% confidence intervals on the estimated mean value.

The differences between the three approaches are apparent in the relative difference statistics, which also show that the GET operator outperforms the monogenic signal and the structure tensor derived orientation values. The differences are due to differences in the representation of the orientation values and are biased towards the GET operator. The reason is the actual differential filters used. The monogenic signal orientation was estimated with an optimal parameter of $\alpha = 1.75$ and $\beta = 0.1$ with the Mellor-Brady filter which has high pass properties therefore the circle edge is picked up as a sharp feature and leads to a dispersion of the orientation values around the circle rim. The ST results may be explained drawing a comparison with respect to the Scharr filter properties. Finally, acting with Gaussian filters on sharp ridges produces better defined features.

The criterion for a good orientation estimator is not necessarily the best accuracy measured on ideal probes, ideally it is also characterised by how well it performs in the presence of noise and with real tissue boundaries. We have found that the monogenic signal performs better on real medical images both in the presence of noise and with respect the predictability of the outcome despite the fact that the synthetic examples presented here do not reflect this.

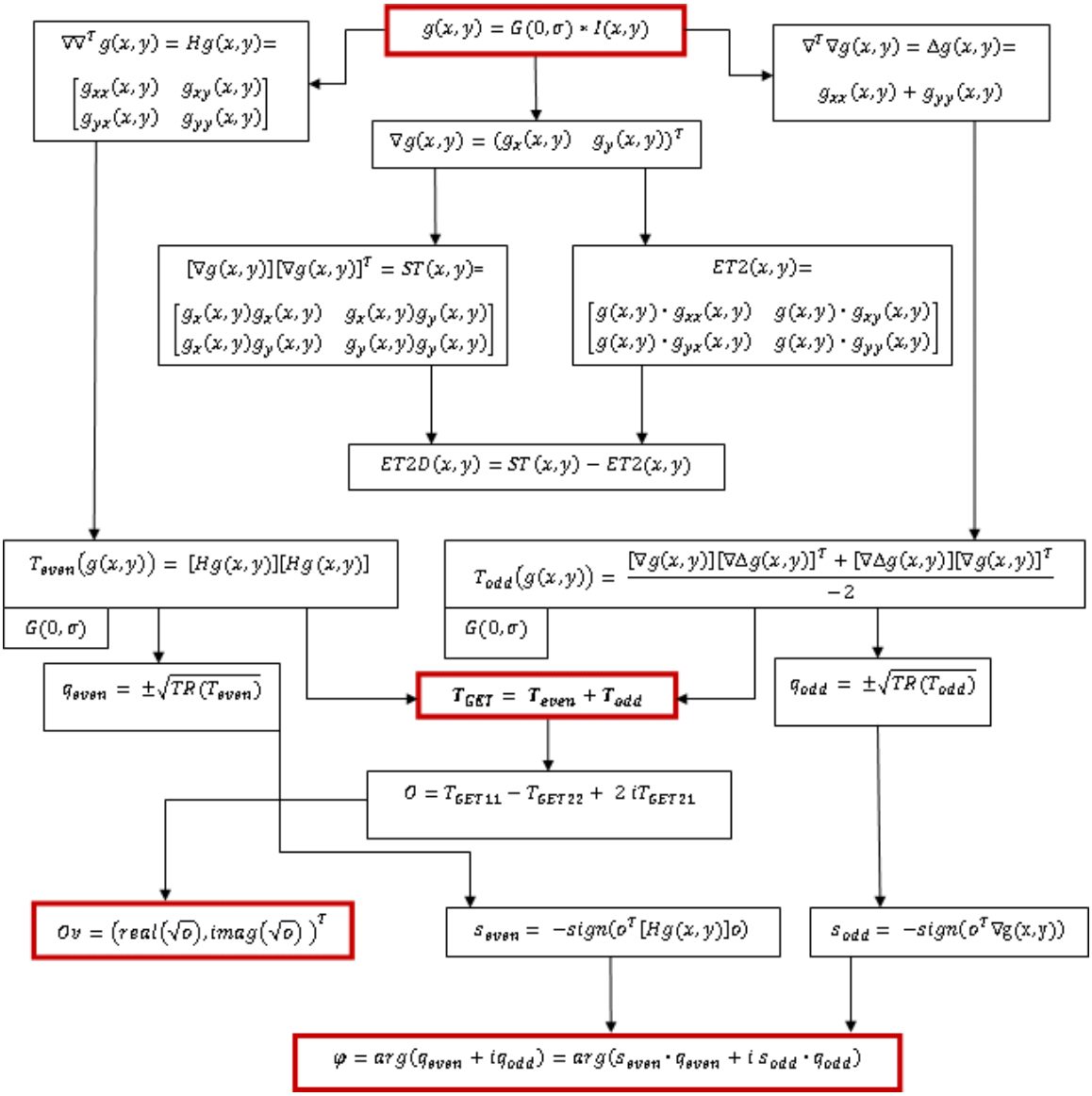


Figure 5.10: Computations behind the Gradient Energy Tensor (GET) schematic, where G is the Gaussian kernel and I is the input image. Notations: $g(x, y)$ is the Gaussian smoothed image, ST denotes structure tensor, $ET2D$ is the 2D energy tensor, H is Hessian, TR is trace, o is orientation vector.

Local energy estimation and statistical description

Local energy in the *ciD* framework is the measure by which we make a decision between feature and non-feature. As described earlier in this thesis all approaches to estimate local energy rely on some form of gradient computation; the *ciD* framework in specific makes use of the squared gradient magnitude image and it was implemented

with central differences using Matlab's built-in function.

5.3.4 Overall system

The steps of the computations include an iterated exchange between the ciD calculation and the iterated optimization within the HMMF (Marroquin et al., 2003) to define the energy and the noise mean and covariance with an assumed mixtures model. While the original model performs well with the assumption of Gaussian intensity tissue class, the successful estimation of the separation between structure and noise strongly depends on the assumed parametric model.

The update of the hidden label field has two additional options included based on Geman and Geman (1984) in form of a line process. Line processes, at least in their original formulation, assumed that there is a secondary label field (here, hidden field) that may be used to constrain the weight between different elements in a neighbourhood. The conceptual meaning behind this step is that no penalty should be given where a line is running between two pixels.

In our implementation, line processes were assessed both as binary decisions based on the ratio between the ciD probabilities (i.e. whether we have high $ci1D$) and line weights in form of the $ci1D$ probability. An illustration between the solution with and without the line process is shown in Figure 5.11. Whether this is desirable is arguable.

A second consideration is the inherent constraints between the uncertainties of the ciD triplet. The propagation of the uncertainties from the $ci0D$ (noise perspective) is included in the uncertainty estimation as described above. The additional constraint due to the geometric uncertainties is present between the $ci1D$ and $ci2D$. Their uncertainty is inversely proportional and it will be bound to the uncertainties of the energy $\left(\overline{E}, \Sigma_{EE}\right)$.

Main computations

The main computational steps and the considerations are included below.

Update steps

We followed the update as follows:

1. Estimate the $\left(\overline{E}, \Sigma_{EE}\right)$ and $\left(\overline{N}, \Sigma_{NN}\right)$ using the HMMF framework with uniform neighbourhood weights. This includes the iteration between the p-field and class update with the assumption of a mixture of exponential.

2. Estimate the uncertainties in the input orientation field.

3. Estimate the ciD triplet and their uncertainties.

Then iterate between:

1. $\left(\hat{p}, \hat{\theta}\right) = \text{argmax}_{p \in S_M^N, \theta} P\left(p, \theta | I\right)$, where $\hat{p}$ and $\hat{\theta}$ denote the estimated hidden class p-field with the parameters for the respective mixture model, and S_M^N is the simplex defined as $S_M = \{u_k \in \mathcal{R}^M : \sum_{k=1}^{M} u_k = 1, u_k \geq 0, k = \overline{1, M}\}$, with M denotes the number of classes and N is the cardinality of the label field. The update on the *p-field* proceeds with an updated weight between the neighbouring elements with respect to the relative distance between the $ci1D$ and the sum of $ci0D$ and $ci2D$ measure.

2. Given the updated energy and noise parameters the ciD triplet is calculated and their uncertainties are estimated.

Computational challenges due to the choice of the exponential model arise when the background model attains a mean/variance value that is close to zero, since this introduces numerical instability. A computational way around this is by the re-normalization and the shift of the actual gradient magnitude values. The actual noise

and structure model parameters are then computed in a second step by the current separation between the structure/noise model based on the hidden label field.

The computation of the potential field was performed using the line-process with the ciD probabilities. There is a two way information update between the estimation of the ciD probabilities and the estimation of the soft-threshold function parameters using the HMMF framework. The ciD probabilities are used as a constraint on the update of the potential field, by selective weighting the cliques based on whether a line is present between two pixels.

Computational choices and their limitations

In our methodology, we took steps to make informed decisions about the choices of the statistical descriptive models, however we acknowledge that both theoretically and computationally there is room for improvement in these.

The current model is based on the existence of a global energy and an isotropic noise model, the validity of both of which is arguable with respect to the ciD framework. The height of the cone is determined by the decision between feature and non-feature, while the radius at which the feature is positioned is due to its geometric properties. However, these are linked by a nonlinear one-way process. The extent to which the geometric uncertainty is modified by the presence of noise depends upon the underlying ciD triplet, therefore the geometric shapes of uncertainties due to noise are not deformed linearly to the same extent.

The major challenge is to describe the local geometric properties of signals with the effect of noise. Due to the lack of a proper mathematical formalism to describe the transition between different iD features these were held back until now. Denoising using local covariances have been already proposed, however without the fundamental constraints from the geometric relations described by the ciD cone. The implementation of this process would include two objectives: (1) estimation of local noise

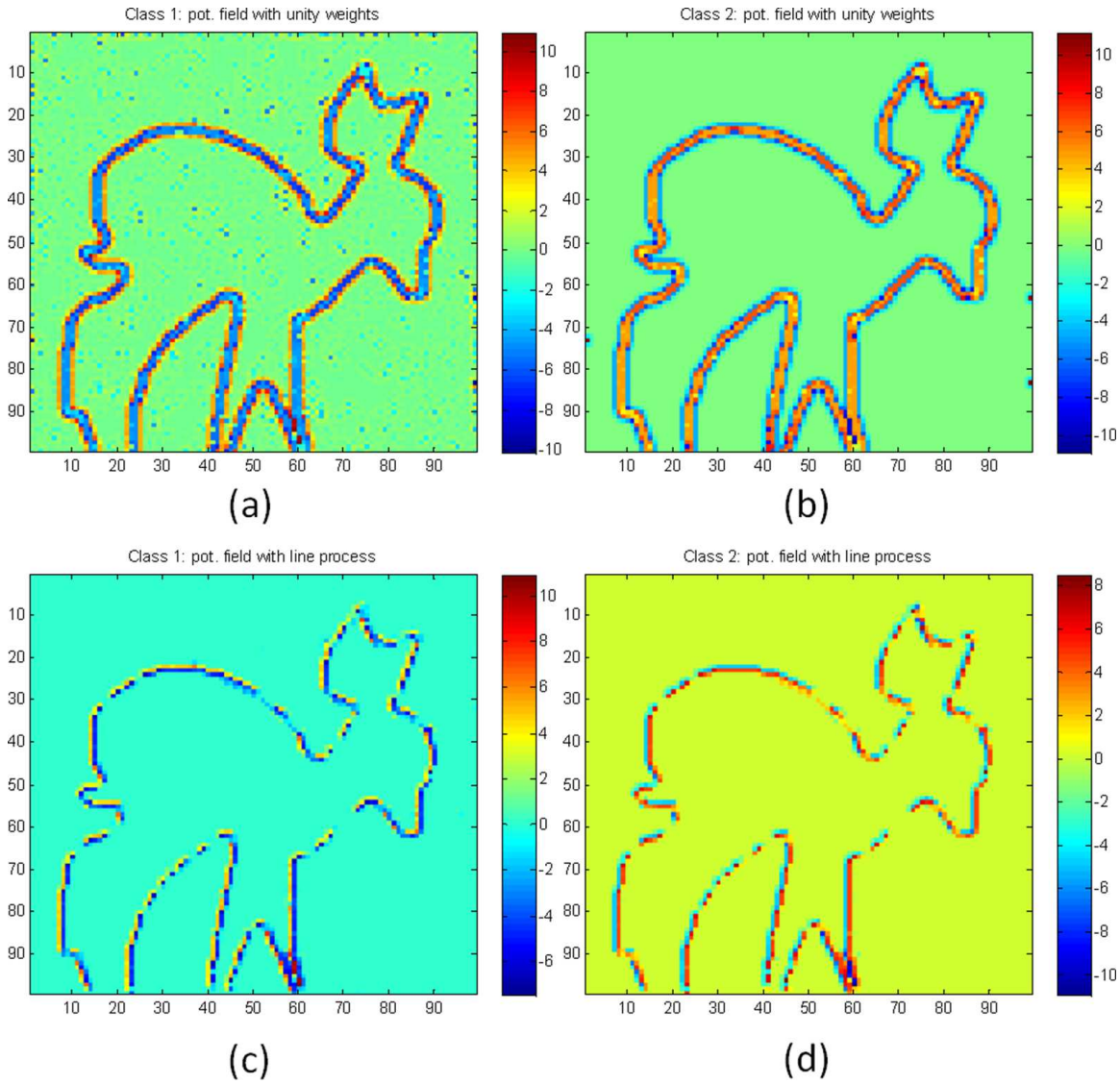


Figure 5.11: Comparison between the potential field for the structure and noise class using the Hidden Markov Measure Field model, with and without line process. The figures (a-b) show the potential field for the noise and structure with unity update weight, and (c-d) shows the potential field with the constraint on the cliques derived from the ciD probabilities.

properties and (2) minimize the ciD field geometric uncertainties.

5.4 Results and Conclusion

Preliminary results of uncertainty estimation in the context of the continuous intrinsic dimensionality are presented in Figures 5.12 and 5.13. Results using natural pictures

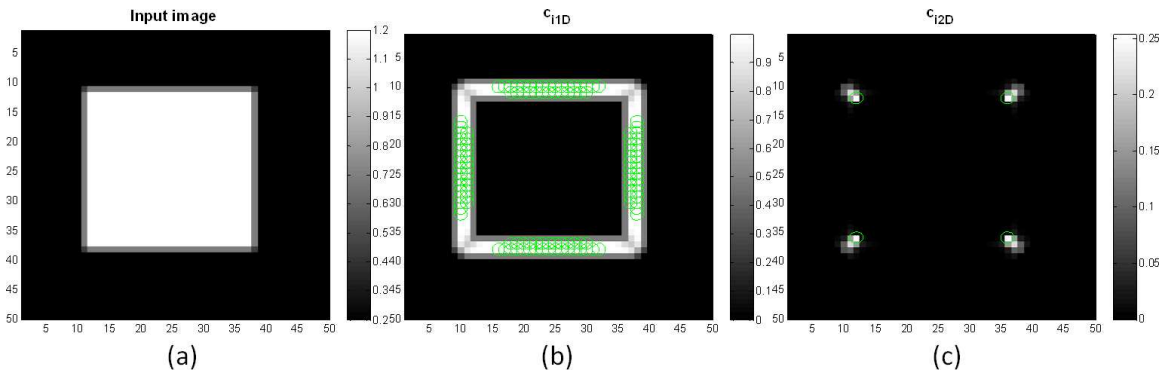


Figure 5.12: Uncertainties labelled with green circles in an ideal rectangle example with 10% noise. Fig. (a) shows the ideal input image with smoothed edges, while (b-c) presents the uncertainties labelled on the ci1D and ci2D probability maps as deduced by the set of computations suggested in the text of this thesis.

and medical images are not included at this stage because the specific choice of feature descriptors used in the methodology has yet to be automated.

Figure 5.12 shows an ideal rectangle image corrupted with 10% Gaussian noise. The derived uncertainties for the line (ci1D) probability and the 2D feature (ci2D) probabilities are indicated by the green circles. The uncertainties were selected using a 40% threshold and the displayed circles scaled to aid visualization. The gap regions where no uncertainty is predicted is due to two main factors: first, it is a low ci1D and ci2D probability region; second, it may be adjusted by the choice of the specific threshold. This figure is only shown as a proof of concept.

A comparison between the salient features detected by the (Kadir and Brady, 2001) algorithm and the uncertainties derived in the ciD framework is illustrated in Figure 5.13. This specific test image was chosen because it includes two types of features and symmetries: the centre of the blobs may be regarded as a point of rotational symmetry, while the rims of blobs denote a line symmetry axis. The scale-saliency algorithm (Fig. (a)) selects the salient features with a joint feature-scale localisation in an isotropic framework. However it does not differentiate between the different types of symmetries. In contrast, the two axis of the ciD triangle allow the detection of the two type of symmetries (Fig. (b-c)). Along the x_t axis the line

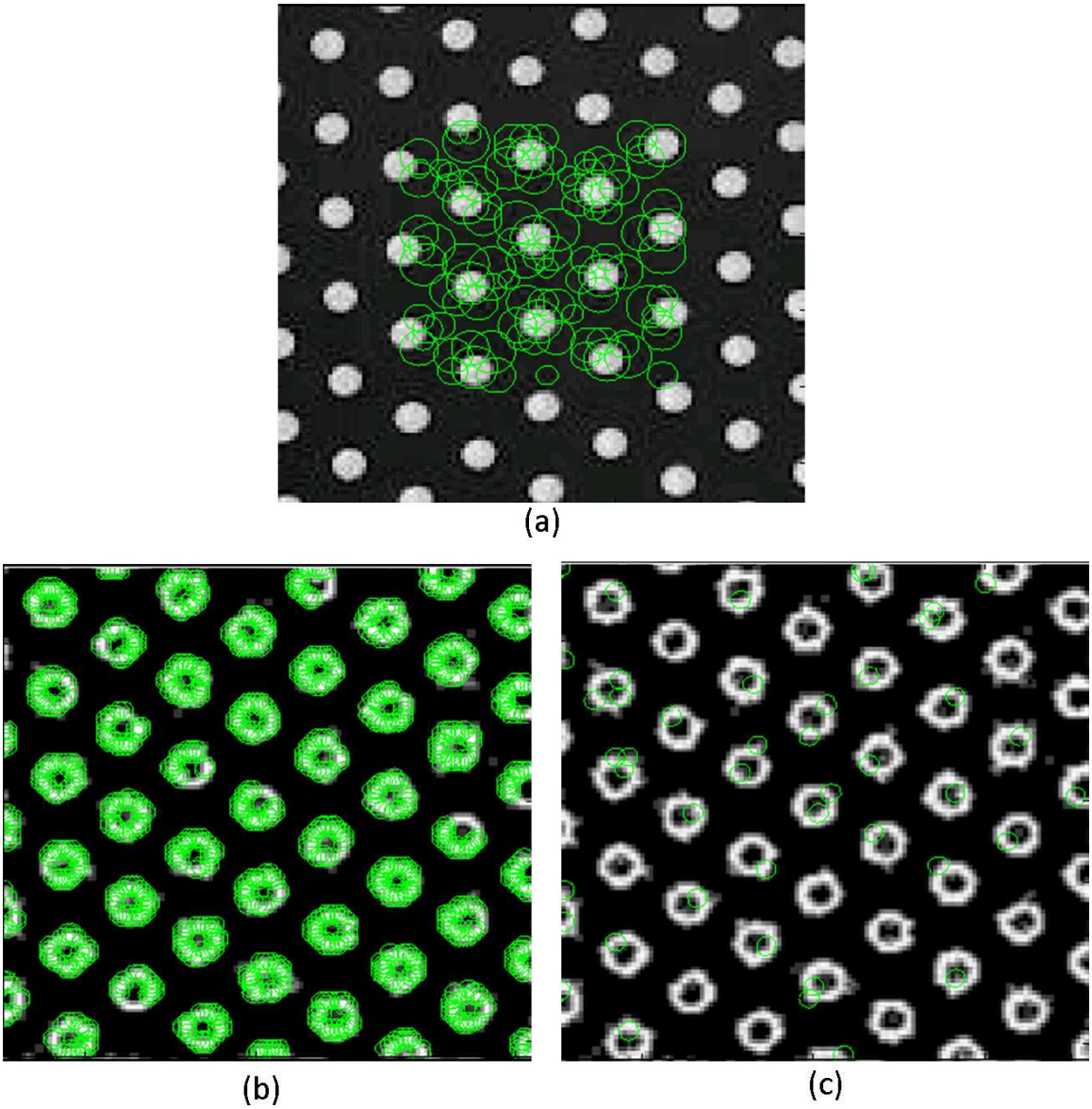


Figure 5.13: COMPARISON: scale saliency map (a) and the uncertainties detected on the edge on the ciD triangle (b-c). Uncertainties are labelled with green circles. The two orthogonal axes of the ciD triangle correspond to the line (b) and origin (c) variance. The detected salient points were chosen with a 40% threshold level. The main differences between the actual results from the two methods are explained in the text and resemble a complementary relationship. For illustration purposes the green circles on the ciD triangle axis map are scaled by a constant factor.

symmetries may be detected as is also suggested by the uncertainty centres derived; while the y_t axis in general would allow the detection of centroid points. Figure (b) shows that the line symmetry centres are marked by uncertainties present, which is

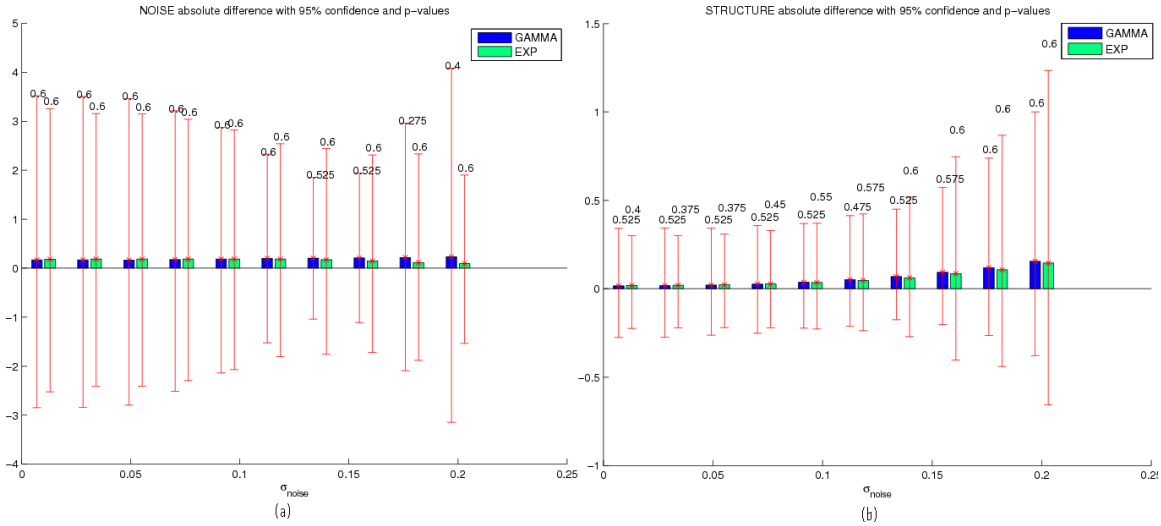


Figure 5.14: Comparison between the descriptive capabilities of the mixture of GAMMA and Exponential models. The method shows the statistics calculated over the absolute difference between the expected an estimated mean from the derived tissue parameters with the 95% confidence intervals and the p-values.

what one would expect. However, Figure (c) shows that there are only sparse ci2D uncertainties detected along the edges of the blob, i.e. it is merely due to discretization error, instead of valid ci2D events. The reason behind this is the selected scale on which the ciD probabilities are detected[3].

In summary, Fig 5.13 suggests that the two approaches are complementary: the scale saliency algorithm achieves its robustness by its approach of joint feature-scale-space localisation, with no specificity regarding symmetries; in turn the ciD methodology (at current) has no intrinsic notion of scale, but it has the potential to differentiate between the different types of symmetry axis.

In a second experiment, we investigated the benefits of the iterative and single step structure and noise parameter estimation, and the possible improvement of the mixture of Gamma distribution model over the mixture of Exponentials. The effective noise and structure parameters as derived from the HMMF output are generally not

[3]With the uncertainty estimation we used we also observed a dependency on the axis orientation, which is a drawback of the estimation methodology as compared to the scale-saliency algorithm proposed in (Kadir and Brady, 2001).

accurate, as was suggested earlier; therefore, based on the classes separated by the label field the parameters of the noise and structure class were derived as a separate step.

The hypothesis that the Gamma function is a possible better representation of the mixture of structure and background tissue properties on squared gradient magnitude images was tested and the results are presented in Fig. 5.14. For these results, images from the KIMIA database (KIMIA-database, 2011) have been used and corrupted with additive Gaussian noise. For each noise level and each structure and noise class in the image, the absolute difference between the estimated and the actual measured mean values were computed. The estimated value was calculated based on the class separation derived from the hidden label field with an MLE fit. The assumption that the absolute mean differences follow the Normal distribution proved to be false and instead the uniform distribution was used. The 95% confidence intervals were derived and the p-values evaluated using the goodness of fit measurement. These are represented in the Fig. 5.14.

The results presented here suggest that the GAMMA distribution leads to an improvement in the description of the structure part until a threshold point defined by the extent of the Normally distributed noise added based on the results presented in Figure 5.5.

The advantages of the iterative update between the ciD and noise structure parameter estimation is shown in Figure 5.15. The figure compares the accuracy of the iterative framework against the single step structure/noise model parameter estimation. The measure of accuracy is computed as the sum-of-squared distance (SSD) between the derived ciD probabilities and the desired ciD probabilities as a function of increasing peak signal-to-noise ratios (PSNR). The results are presented both as a function of noise variance (first row) and as a function of the PSNR to aid a better comparison. The two parallel line example suggest that a systematic better accuracy

with the iterative framework can be achieved as opposed to a one step signal to noise parameter estimation (as is suggested by the original model formulation).

In conclusion, several advances to the current formulation of the ciD have been proposed. We suggested that the exponential distribution is not particularly suited to model structure for general and medical images. Instead the Gamma distribution showed very promising results. A Hidden Markov Measure Field based method was also proposed as well as an alternative way to estimate the ciD uncertainties, and the concept was proven on ideal feature examples. Theoretical models were proposed with geometric constraints for the ciD uncertainties, which however were not implemented in this thesis.

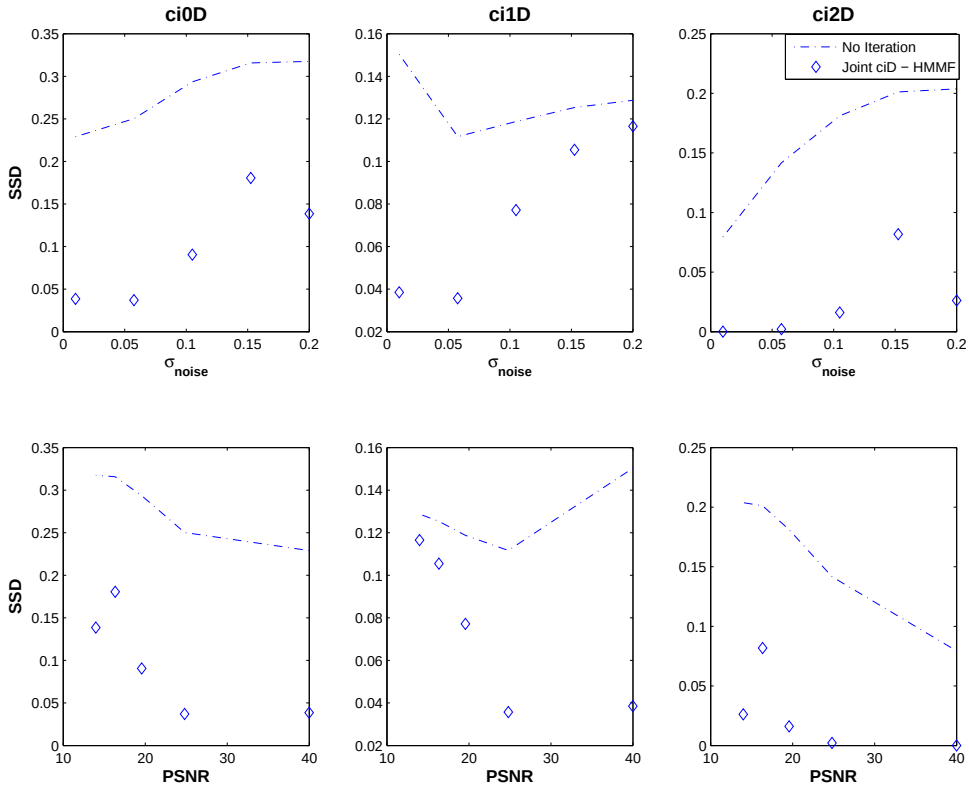


Figure 5.15: Comparison between the continuous intrinsic dimensionality estimation and the estimation of the noise versus structure model parameters using the Hidden Markov Measure Field framework.

Chapter 6

Validation: Liver MRI tissue characterization using phase based measurements

In this chapter we review the steps we have made to quantify the texture in T_1w liver MRI as a surrogate marker for liver fibrosis. We do this because currently there are no methods that allow one to derive such measures non-invasively. We proceed by describing our data set, the pre-processing of the images that is necessitated by the bias field effect, and the computation of the texture measures we computed for each patient image. This last step uses as underlying texture measure the Riesz weighted phase congruency (as described in Chapter 4), and the spatial summary of this measure by using the fractal dimension along with the comparison with and without correction for the bias field effect. In addition to the fractal measure, we compute an entropy estimated using NP windows on the closest distance maps of the derived features and classify them using k-means. We also compute a "clumpedness" measurement from (Clark and Evans, 1954) and, for completeness, the Laws' texture measure.

There are two main goals of this chapter:

1. Differentiate by trend between early, mild and severe fibrosis, i.e. Ishak scores 0-1, 2-4 and 5-6 respectively.

2. If possible, make a separation between pericellular and non-pericellular fibrosis. This is a special type of fibrosis which is currently often missed and for which there is no means of separation except biopsy.

There are several results that we show below:

1. Liver tissue is a rich source of textural features. For this reason we have applied a number of computational methods, for example fractal based measurements (Lee et al., 2003), to map textures, hence to quantify disease stage. We found that the fractal-based measurements provide a reasonable differentiation between the various Ishak scores with respect to median trends. In addition, fractal-based measurements differentiate well pericellular and non-pericellular fibrosis within the same Ishak score with the bias corrected phase based measurements.

2. Motivated by the concept of local binary patterns (Wang and He, 1990) we derived local entropy map distances. Our conclusion based on the experiments were that the closest distance-based entropy maps did not correlate well with the Ishak score trend, but they do not provide useful information to differentiate between pericellular and non-pericellular fibrosis.

3. Based on the observation that fibrosis appears in spatially restricted regions we have evaluated the use of clumpedness measurements borrowed from population spatial statistics. The clumpedness measurement does not results in the expected trend between the different Ishak scores for all three cases as detailed above.

4. In conventional texture analysis, the use of Laws filters is a standard method that is also used in medical imaging to quantify texture differences. The average of the derived Laws energies within a region of interest was used, which showed sensible statistical correlation with respect to Ishak trends. There are two representations shown to aid understanding.

All the above observations were derived based on the Riesz weighted phase congruence (PC) suggested in the earlier chapter. All the results were compared to the standard Kovesi type PC, which gives a poorer results. The best method proved to be the Laws average values calculated based on the Riesz weighted PC.

We first describe the data set and the corrections applied, such as bias field correction. The bias in magnetic resonance based data recording severely limits automatic analysis. It is a result of B_1 (Radio Frequency) field inhomogeneities, which are not only influenced by the receiver coil geometry and gain properties but also by the anatomy, that is by the loading of the system. The resulting intensity images are characterized by a slowly varying ramp in intensity. It is the low bias field that is of particular interest to us, since this results in signal loss, low contrast, limiting intensity based segmentation, based both on regions and edges.

The preprocessing is inevitable, given that the bias field affects not only the intensity, but also the local phase based measurements. Although it does not seem to have been reported previously, we observe that the local phase based liver maps are extremely sensitive to this signal drop, which persists over multiple scales. As a result, the phase congruency type feature model detects it as important structural edge/ridge information.

Unfortunately, most developed techniques are not directly relevant to our specific images, as explained in the following section.

6.1 Liver images and the effect of the biasfield

The liver MRI was provided by Dr. Rajarshi Banerjee, Churchill Hospital, Oxford. All the liver images are part of a liver study (RIAL - Rapid Image Assessment of the Liver) in which there are three sets of data generated: a T_1 map, T_2 map and proton spectroscopy to assess liver steatosis. In the original study the quantitative liver T_1 values are used, recorded with the so called Shortened Modified Look-Looker Inversion recovery pulse (ShMOLLI) (Dall'Armellina et al., 2012). It turns out that the iron content of the liver has a material impact on the T_1 estimates, and this can be detected using T_2 maps, as is performed in the original study. Biopsy readings are used as the ultimate diagnosis tool and ground reference for image interpretation. Given that any rating depends on a human observer this is only an approximate ground truth for validation. However, this is currently the best available.

The liver images are recorded during an approximately 20 min scan with 15 slices to guide the position of the T_1 mapping within the liver. The in plane resolution of the images is 1.19mm $\times$1.19mm with 6mm gaps. These images are T_1w fat-SAT images (fat suppressed, which means that the fat tissue signal is saturated so that no bright regions are shown). However, these have not been optimized for contrast, motion reduction, or other common MR associated artefacts. This is because the actual images are not used for the examination of the disease state of the patient's liver.

Despite their suboptimal contrast and poor resolution there are good reasons to work with this data set. First, there is a statistically significant number of liver data available, with associated Ishak scores, which is absolutely necessary for the validation of the developed methods. Second, we have access to the expertise of the clinicians with respect the interpretation of the images. The T_1 maps are not available in standard Dicom format and it is not straightforward to interpret them. In parallel with this, the T_2 maps are only useful for the quantification of the iron

content. Therefore, our analysis contains the T_1w fat-SAT images that are closest to the standard liver MRI.

To date 19 patient data sets have been made available to us. However, only 15 of them have an associated T_2w images that is used for vessel extraction and the experiments reported below will be only presented for these.

The area of the liver has severe degradation due to the bias field effect (see Figure 6.1). As noted, this inhomogeneity always appears at approximately the same location: the upper and left hand side for each patient with signal drop, as observed on the axial scans. Several MRI bias field correction algorithms have been proposed, with very good results reported for the brain and spine. There are some bias field correction methodologies reported for the liver, but these usually rely on the specific pulse sequence and the specific images recorded. There is no gold standard available to be used on the liver due to its complex tissue texture.

Although it is not our main focus, we cannot exclude the bias field correction step in our analysis. The bias field appears as a slow change in intensity images, which precludes most intensity based automatic image analysis. Unfortunately, all phase congruency feature models, including the one that we have developed, detect these regions as persistent changes, leading to false detection of liver structure.

Our data has several limitations, not just in terms of needing bias field correction. In general, brain scans are performed using a birdcage coil that is tightly fitted around the patient's head, leading to exceptional signal-to-noise ratio (SNR) and contrast, and minimal, if any, motion artefact. On the contrary, our liver MRI were recorded with a surface phased array that lies on the patient's stomach. Given that the abdomens of patients vary tremendously, so does the effect of the bias field. In addition, the computational model of the brain may be modelled as being piecewise constant, whereas the liver contains a significant number of structure changes that are informative of pathology and as a result should not be suppressed.

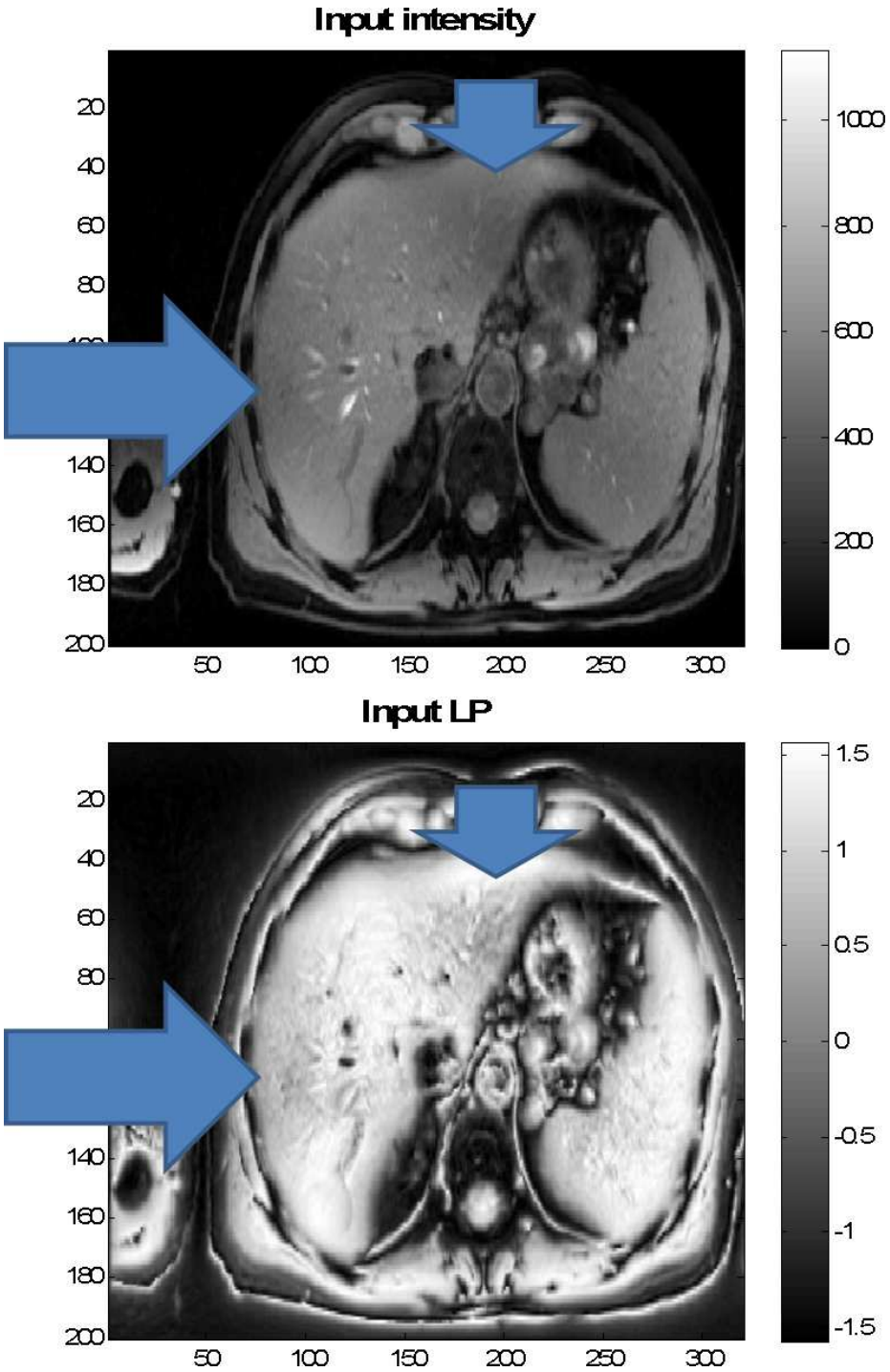


Figure 6.1: Bias field effect illustration: TOP intensity T1w image, BOTTOM local phase map on a single scale generated using the monogenic signal's framework.

6.2 Pre-processing the liver MRI: biasfield correction

Several methods have been reported for MR bias field correction. The earliest methods used homogeneous phantom measurements to derive the bias field. A major drawback of this approach is that the influence of the loading of the system is not taken in account and to some extent may vary from patient to patient. There are also physics based approaches that take in account the basic MR signal formation to correct the images. For example, see (Noterdaeme and Brady, 2008), where the bias field estimation is performed using calibration images recorded at multiple flip angles. Finally, there are computational approaches that operate directly on the images (Styner et al., 2000), (Wells et al., 1996), (Guillemaud and Brady, 1997). These are based either on global or local analysis of the images.

(Styner et al., 2000) propose the PABIC (PArametric BIas field Correction) correction with a polynomial function to describe the smoothly varying bias field. Specifically, it uses higher order Legendre polynomials to make a close approximation to the bias field. The parameters of the polynomials are then estimated using a nonlinear optimization. The method has proved to perform well on brain images, however it has enjoyed only limited success on breast MRI. This is possibly because of the base assumption of the piecewise constant tissue classes, which is no longer satisfied.

An advance in the form of simultaneous bias field estimation and tissue classification is presented in (Wells et al., 1996) and (Guillemaud and Brady, 1997). This uses an Expectation-Maximization (EM) framework to iterate between tissue class probabilities and bias field update. The limitation of the methodology is that a good estimate of the tissue parameters described by a Gaussian function are required, and it still relies on the assumption of piecewise constant tissues. The (Guillemaud and Brady, 1997) method is an extension of (Wells et al., 1996), in which a 'sink' tissue

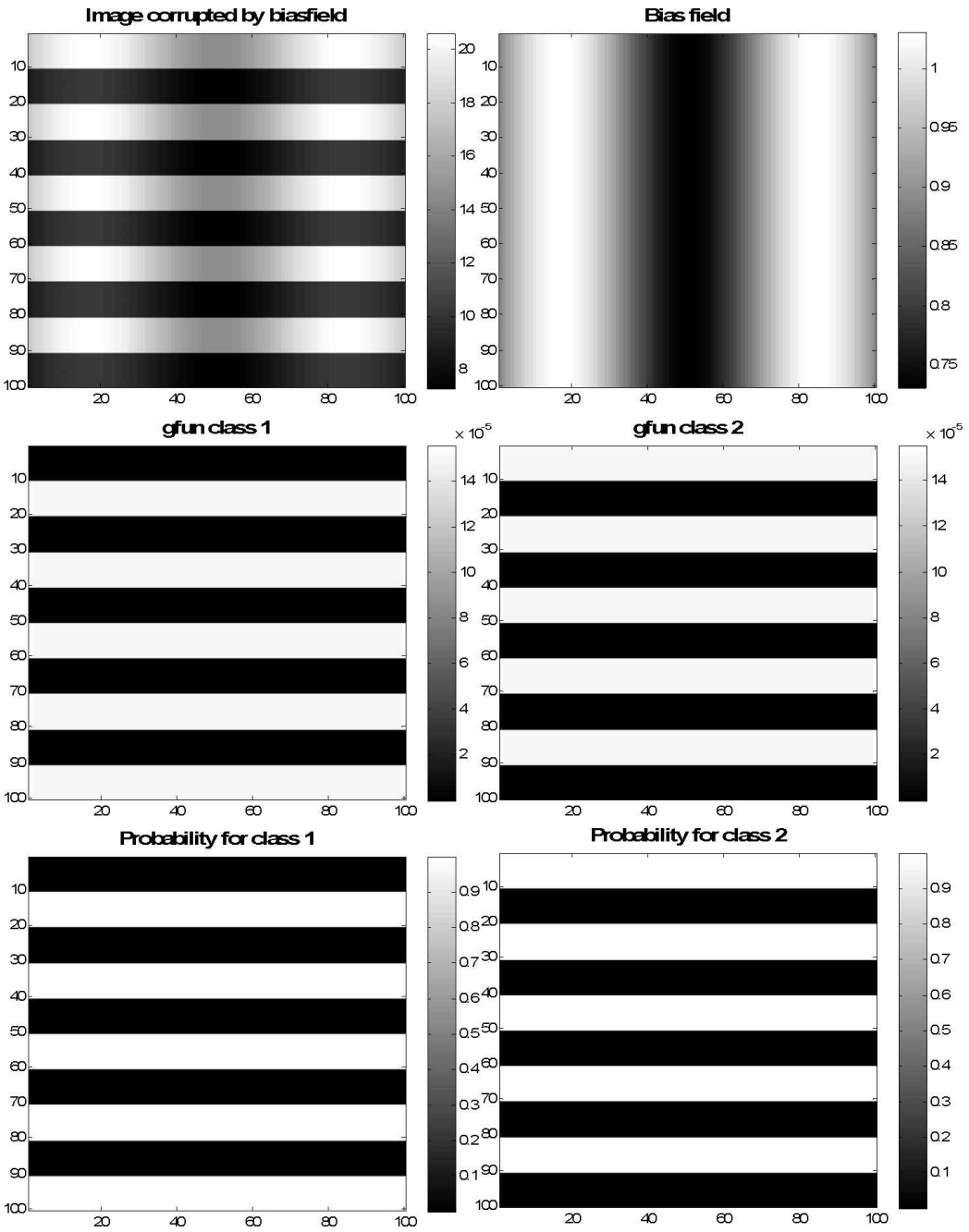


Figure 6.2: Bias field removal from synthetic images. The maps displayed are to aid interpretation of the described methodology steps. The images in the first row show the bias field corrupted two class image and the ground truth sinusoid bias field. The second and third rows show the evaluated Gfun and tissue class probabilities (W) for the two classes at the final iteration. For more information, see text.

class is included and which is assumed to follow a uniform distribution. This was introduced to allow the classification of all the rest of the organs apart from the one

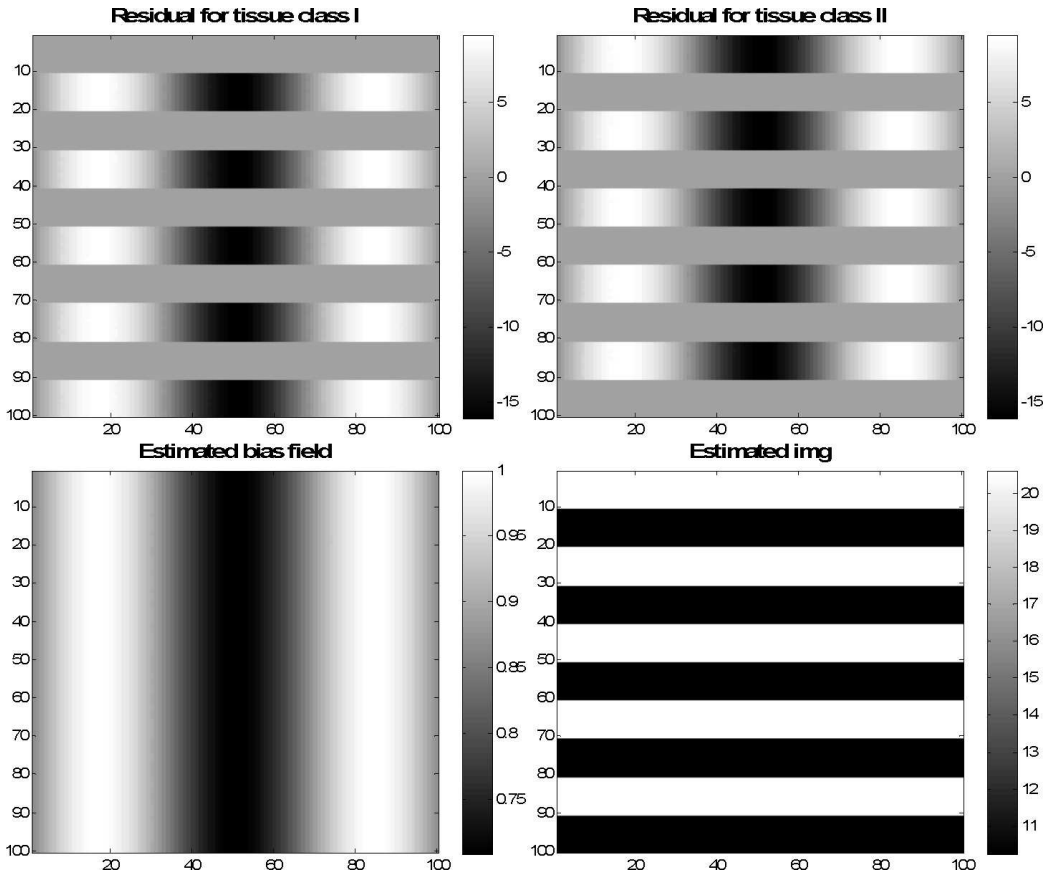


Figure 6.3: Bias field removal from synthetic images. The maps displayed are to aid interpretation of the described methodology steps and is a continuation of Figure 6.2. The top two figures show the residuals for the two tissue classes, while the final row includes the estimated bias field and the corrected image.

of interest.

The methods mentioned so far rely on the intensities. However, in our case, we observed a substantial change in the local phase based maps of the liver and surrounding anatomy, as presented earlier. Without an understanding of what the differences are in the representation of liver texture, pathology and bias field, it is hard to devise a methodology based purely on local phase. Our texture quantification method is based on phase alignment measured at multiple scales, and does not differentiate between these three classes. These limitations may be alleviated by studying the phase representation of the three classes from images annotated by clinicians; this would be an interesting future path to pursue.

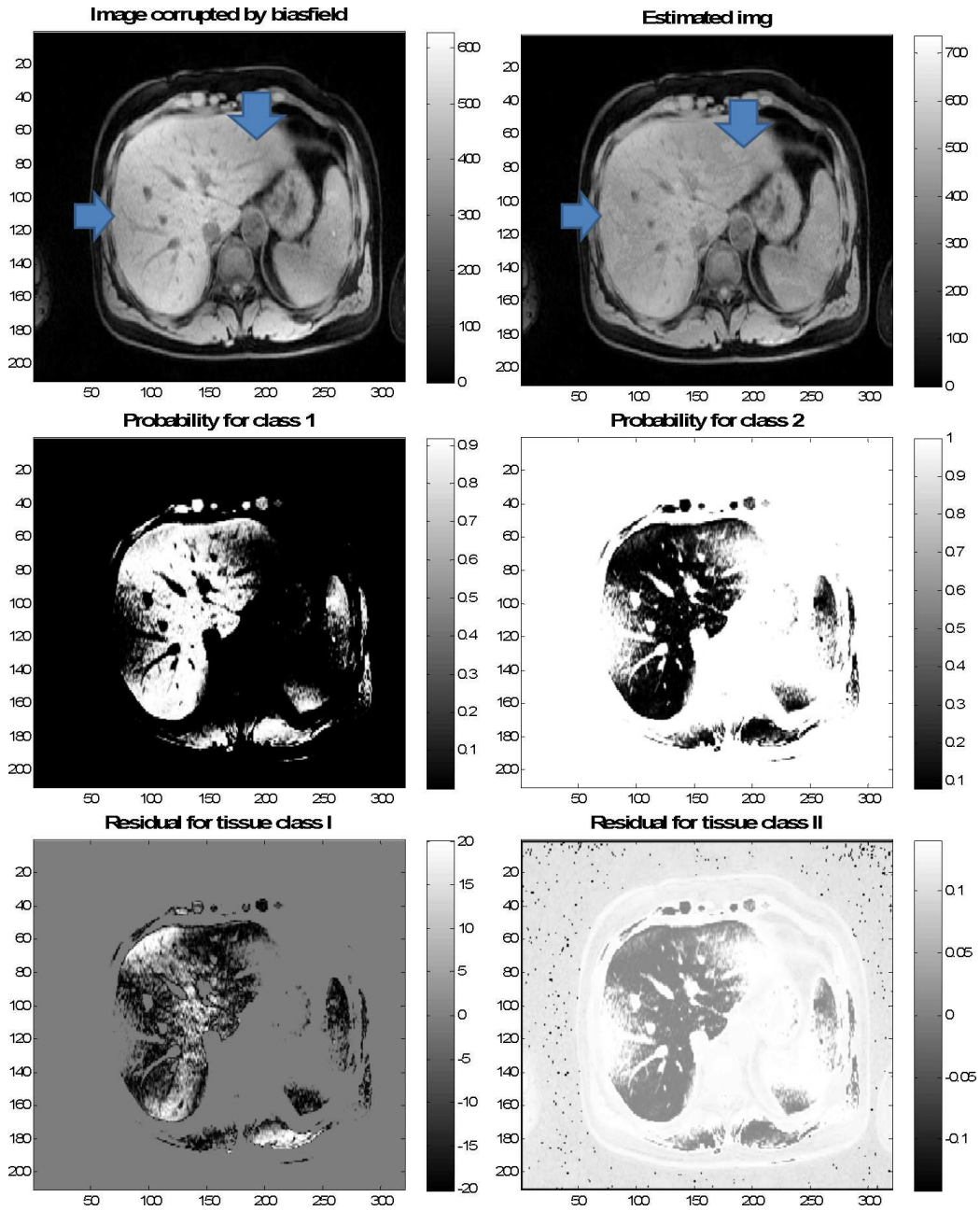


Figure 6.4: Bias field removal from liver MRI. These are included to aid understanding of the bias correction algorithm steps. The first row contains the bias corrupted and corrected images, where the arrows point to the affected region. The class probabilities (W) are shown in the second row, while the residuals are in the final row. For more information, see text.

Given that we are only interested in bias field estimation as a pre-processing step, and given the above mentioned challenges with phase-based bias field representation, we concentrate on intensity changes. More specifically we follow the methodology

presented in (Wells et al., 1996) and (Guillemaud and Brady, 1997), especially the latter.

In MRI, the bias field is assumed to have a multiplicative effect, simplified to:

$$I_{\text{measured}} = I \star \mathcal{B} + \mathcal{N} \tag{6.1}$$

where I_{measured} is the observed bias field corrupted image, I is the original image, $\mathcal{B}$ denotes the bias field and $\mathcal{N}$ is additive noise. The $\star$ operator denotes point wise multiplication. Given the above multiplicative form, the calculations are performed in the logarithmic domain, where the dependencies can more easily be separated.

The tissue models we used are analogous to those presented in (Guillemaud and Brady, 1997). Despite the fact that the piecewise constant model might not be entirely appropriate for our images, as we shall show, it performs reasonably well to correct the phase maps. It also leads to differences in the final texture based fibrosis scores we are interested in. As noted in the original paper, the method relies on a good tissue parameter estimate. While this may be possible for the brain, it is infeasible for the liver given the high variations encountered with the non-optimized sequences.

As a result, we have developed a semi-automatic method. In our model, there are two tissue classes, as well as the bias field. The initial mean and covariance of the liver is estimated empirically from a small interactively selected region, preferably from areas not affected by the bias field as (where $\mathcal{B} = 1$ or $\log I_{\text{measured}} = \log I$):

$$\begin{aligned} \mu_{\text{liver}} &= \frac{1}{N} \sum_{i=1}^{N} \log \left(I_{\text{measured}}(i) + \epsilon \right) \\ \psi_{\text{liver}} &= \frac{1}{N} \sum_{i=1}^{N} \left(\log \left(I_{\text{measured}}(i) + \epsilon \right) - \mu_{\text{liver}} \right)^2 \end{aligned} \tag{6.2}$$

where N is the total number of pixels in the selected window and i only runs over the selected pixel intensities, ϵ is a small number within the machine precision of

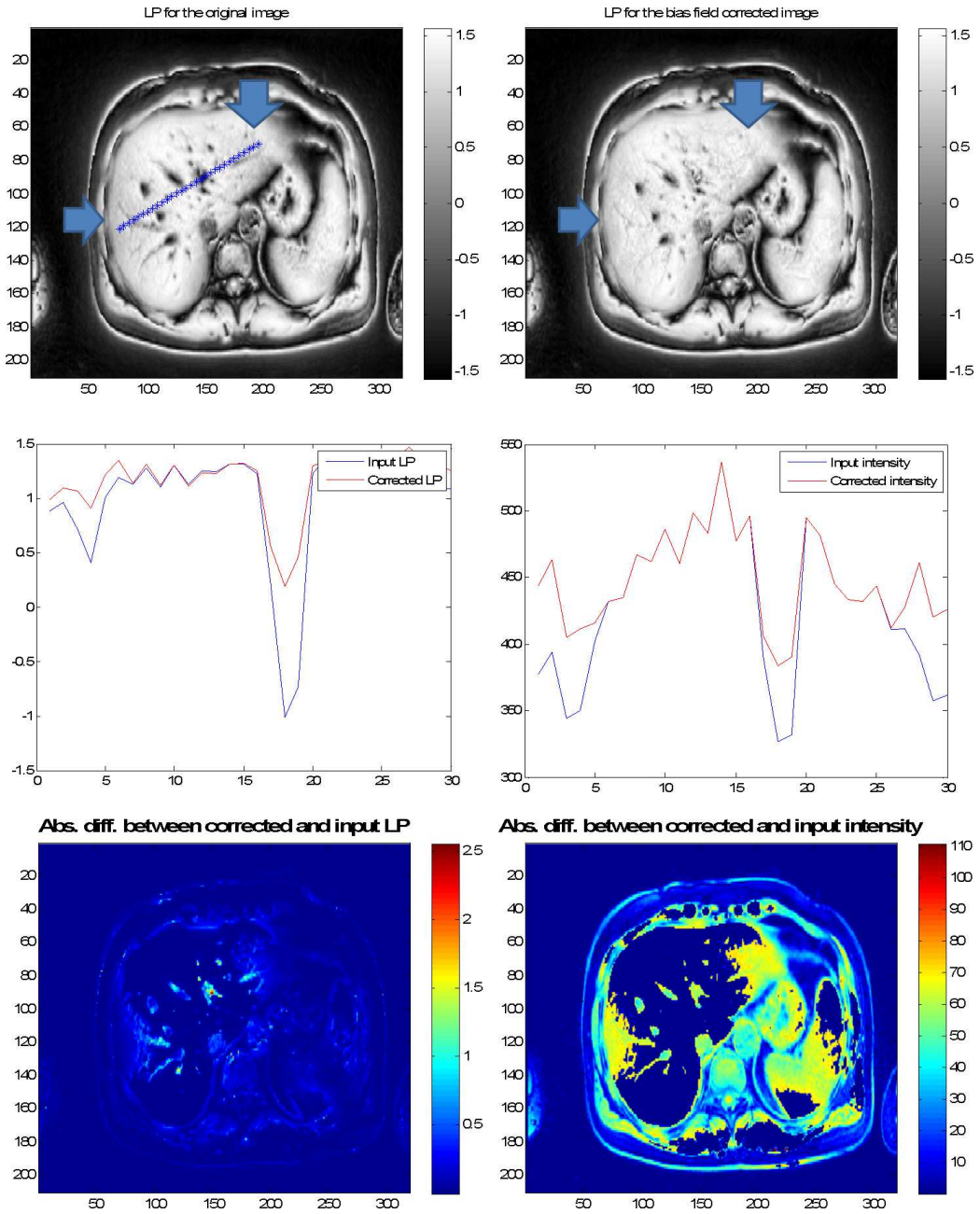


Figure 6.5: Bias field removal from liver MRI. This figure is a continuation of Figure 6.4. The top figure illustrates the bias field corrupted and the corrected phase maps, the arrows point to the affected regions. The second row contains the line profile for corrupted and corrected phase and intensity values along the position presented by the blue line on the first map. The final two images are the absolute difference between the corrupted and corrected phase and intensity maps.

Matlab to avoid taking the log of zero. During the EM iterations the selected window is restricted to regions where the liver tissue class probabilities are greater than or

equal to 0.5. The parameters of the uniform distribution, that represent the rest of tissue in the field-of-view (FOV), are estimated from all the pixels outside the selected window.

This is a workable path for our case, because, as we observed earlier, the bias field mainly affects two to three regions in the liver. The parameters of the liver tissue are fixed during each EM iteration, while the mean and variance of the 'other' tissue changes with each iteration, based on the class probabilities. For example, if the liver tissue probability at position $\overline{x}$ is less than 0.5 then it is assumed to belong to the 'other class' and the empirical parameters for this are calculated. In what follows, to keep notation less cluttered, the spatial coordinates are excluded from the presented formulae.

Compared to earlier methods, we do not initialize the bias field with any prior, despite the fact that based on our observation the regions appear at approximately the same region.

Let $U = [\mu_{liver}\ \mu_{other}]$ be the mean vector and $\psi = [\psi_{liver}\ \psi_{other}]$ the covariance matrix. As a first step, we estimate the deviation from the tissue class mean values:

$$X(\text{tissue}_k) = -0.5(Y - U(\text{tissue}_k) - \mathcal{B}) \star (Y - U(\text{tissue}_k) - \mathcal{B}) \star \psi(\text{tissue}_k)^{-1} \quad (6.3)$$

where $Y = \log\left(I_{\text{measured}} + \epsilon\right)$ and $\mathcal{B}$ denotes the bias field. The class probability for the Gaussian tissue is calculated as:

$$\text{Gfun}(\text{tissue}_k) = \frac{1}{N}(2\pi)^{-1}\psi(\text{tissue}_k)^{-0.5}e^{X(\text{tissue}_k)} \quad (6.4)$$

for the uniform distribution, for all values that are within the range is:

$$\text{Ufun}(\text{tissue}_k) = \frac{1}{\max(\text{tissue}_k) - \min(\text{tissue}_k)} \quad (6.5)$$

Finally, as part of the E-step, the probabilities for each pixel are calculated as:

$$W(\text{tissue}_{\text{liver}}) = \frac{\text{Gfun}(\text{tissue}_{\text{liver}})}{\sum\sum \text{Gfun}(\text{tissue}_{\text{liver}})}$$
$$W(\text{tissue}_{\text{other}}) = \frac{\text{Ufun}(\text{tissue}_{\text{other}})}{\sum\sum \text{Ufun}(\text{tissue}_{\text{other}})} \tag{6.6}$$

where we have a double summation that runs over the 2D parameters. Then the class probabilities are normalized as:

$$W(\text{tissue}_k) = \frac{W(\text{tissue}_k)}{\text{norm}} \tag{6.7}$$

where norm is the sum of the probabilities along the third dimension, $\text{norm}(x, y) = \sum_{i=1}^{K} W(x, y, i)$, where the x, y are the running parameters over the horizontal and vertical direction and K is the number of tissue classes, i.e. the norm is a matrix and we have a pointwise division in the formulae above.

In the M-step we estimate the bias field using the above derived class probabilities and tissue parameters. The residual is calculated as:

$$R(\text{tissue}_k) = \psi(\text{tissue}_k)^{-1} W(\text{tissue}_k)(Y - U(\text{tissue}_k)) \tag{6.8}$$

and for each class

$$\beta(\text{tissue}_k) = \frac{R(\text{tissue}_k) \bigotimes H}{\psi(\text{tissue}_k) I \bigotimes H + \epsilon} \tag{6.9}$$

with H being a circular lowpass averaging filter of unit radius applied along the two spatial coordinates, $\bigotimes$ denotes spatial convolution, and the final estimate of the bias field after cropping the margins due to the convolution and summing over the third

dimension (along the class number)

$$\mathcal{B} = 1 - \frac{\beta}{\max(\beta)} \tag{6.10}$$

Unfortunately, this estimate terminates with a value in the interval $[0, 1]$ while the lower threshold means that we suppress all the signal. Using a synthetic image with two tissue classes, the bias field was estimated without any further processing.

Given that we have the same MR and coil system, the bias field effect varies with the geometry and the loading by the patient. To evaluate the possible reasonable changes in the bias field estimate, an exhaustive search for a reasonable parameter was made within the range $[0.5, 1]$. We found that the classification of liver tissue based on the intensity values was reasonable between $[0.85, 1]$, so all values below this value are discarded and replaced by the lower threshold. This is why our bias field estimates are not as smooth as expected.

Given $\mathcal{B}$, the uncorrupted image is then estimated either by taking the exponential of the difference between the Y and bias estimate or by taking the quotient of the I_{measured} and the exponential of $\mathcal{B}$. We used the second approach.

The steps of the above estimation algorithm are presented in Figure 6.2 and 6.3 for synthetic images; and in Figure 6.4 and 6.5 for example liver image. Figure 6.2-6.3 shows a two tissue texture with a sinusoidal bias field, along with the estimated probability density values and the class probabilities. In this simple case, we get an almost exact approximation of the bias field and the original texture without a constraint on the bias field - as is done for the liver.

Figure 6.4-6.5 shows the same steps, except that this time no ground truth is available either for the uncorrupted liver image or for the bias field. The corrected intensity is not fully homogeneous, as in the synthetic example, but it does correct to a large extent the phase values. The profiles in the intensity and phase values before

and after correction are shown alongside the absolute differences between the 2D maps. Note that the magnitude of the absolute difference is relative to the dynamic range of the actual pixel values. Although the correction changes are higher for the intensity, the phase values are more homogeneous as opposed to the intensity values.

One possible drawback of the method is that it often diminishes contrast between blood vessels and liver parenchyma. This is also observable on the phase images. By examination of the fractal dimension values, however, it is seen that bias correction leads to higher complexity representation. To what extent this is valid for the current texture quantification is dictated by the Ishak scores to be compared.

6.3 Liver image pre-processing: vessel removal

There is a two way argument whether to remove or leave in place the blood vessels and ducts during texture analysis. We had earlier access to a limited data set and this taught us several things. Gadolinium-enhanced high resolution liver optimized images show differences in the blood vessel appearance with the change in the disease stage, therefore they are part of the liver texture. On the other hand blood vessels are not intrinsically part of the liver texture of interest therefore should be excluded from the liver itself.

Vessel removal proceeds based on the realization that the T_2 images contain most of the liver vasculature information. This is because blood vessels and ducts are fluid filled, so they have a low T_2. Unfortunately not all the available images have an associated T_2 image that could be used for this purpose. To extract the same blood vessel in the T_1 image, we need to register the T_1 and T_2 images. We first experimented with two types of registration: the iterative closest point algorithm available in Matlab; and those included in the toolbox MIRT that contains affine registration with three different criteria. Unfortunately, all of these algorithms generated registration results

in which the blood vessels and the curvilinear structures were blurred, so we devised a simpler method. This includes manual adjustment of the axial cross section with the liver in the FOV and doing a bilinear interpolation between the T_1w and T_2w images.

The blood vessels were detected using the energy weighted phase congruency from the $T2$w images using the Log-Gabor filter and a threshold of 0.3, which was manually adjusted. The masks of the liver were manually segmented and the phase congruency detected blood vessels and duct overlaid.

Four example vessel masks are shown in Figures 6.6 and 6.7. The samples correspond to patients RIAL20, 16, 05 and 18. Both images show the T_1w, T_2w images and the generated masks. Each sample is informative because none of the original images reveal the vessel structures detected in the T_2w image.

RIAL20 shows that the original image is very smooth and does not even suggest the presence of blood vessels, while there are clear blood vessels in position (x, y) = (120, 70). One might argue that the detected structure is fibrosis. However, without a biopsy specimen there is no way to settle this question.

RIAL16 has two prominent blood vessels, both around x equals 100 and y between 100 and 150. These blood vessels are visible on both MRI images and clearly segmented in the mask image at exactly the same location. RIAL05 has a clear vessel at (x, y)=(100, 100) position. This is again clearly detected by the derived mask. This image is particularly hard to analyze by non-phase based measurements because the top of the liver contains either fluid or movement and appears particularly bright on the T_2w image.

Finally, RIAL18 T_1w image shows a smooth appearance without possible blood vessel effects. As opposed to this the T_2w image shows a clear tubular structure at (x, y) = (100, 70) position. This branching vasculature is visibly represented in the derived mask.

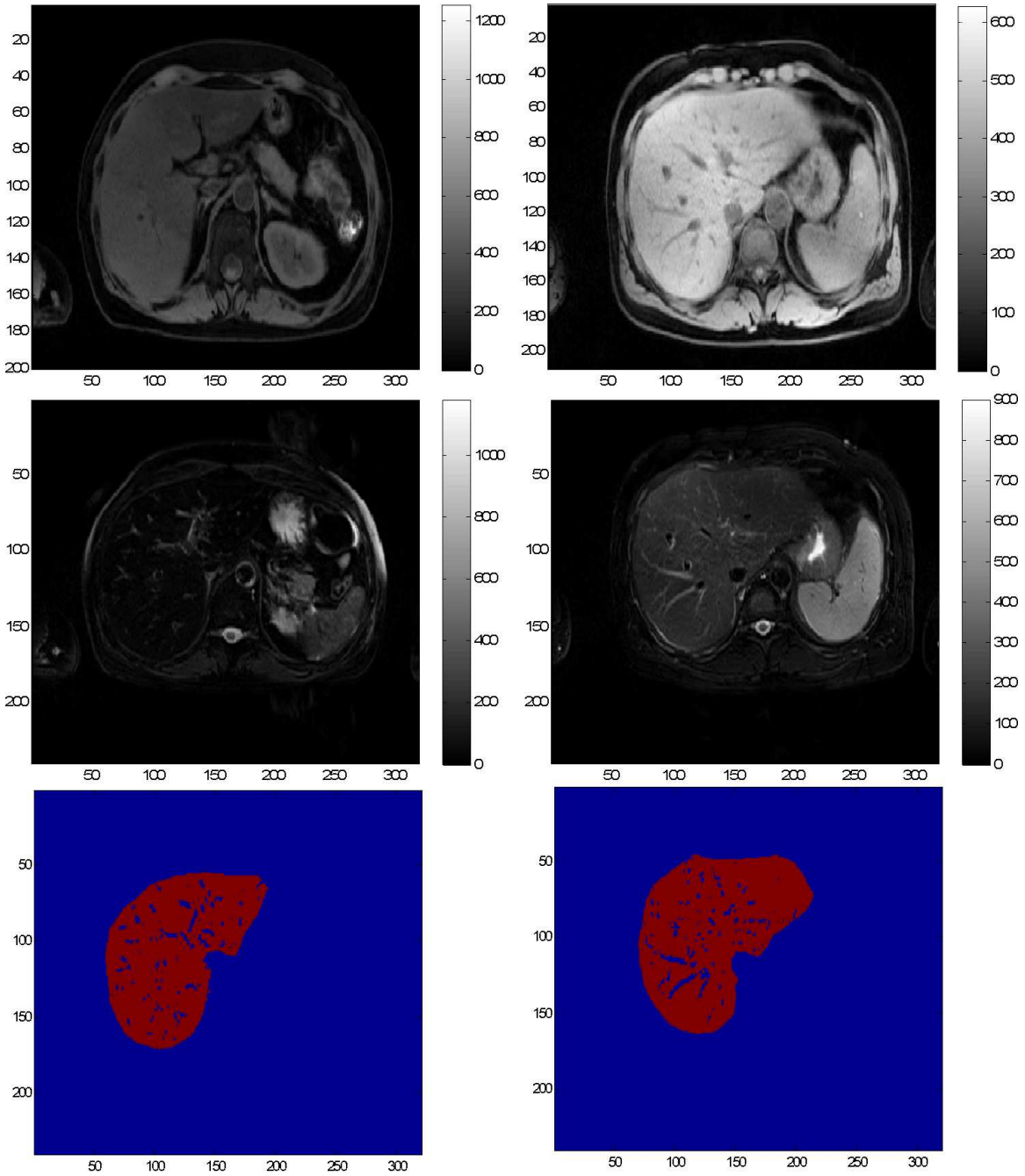


Figure 6.6: Vessel detection from T_2w images. The two columns correspond to RIAL20 and RIAL16. The first row is the T_1w image, the second row is the corresponding T_2w image, while the final row is the detected vascular and duct structures. Note the difference in vascular representation on the two different weighted images in favour of the T_2w. Note the definite vessel structure around x and y axis 150, 100, and 100, 100 respectively and their representation on the vessel detected masks in the final row.

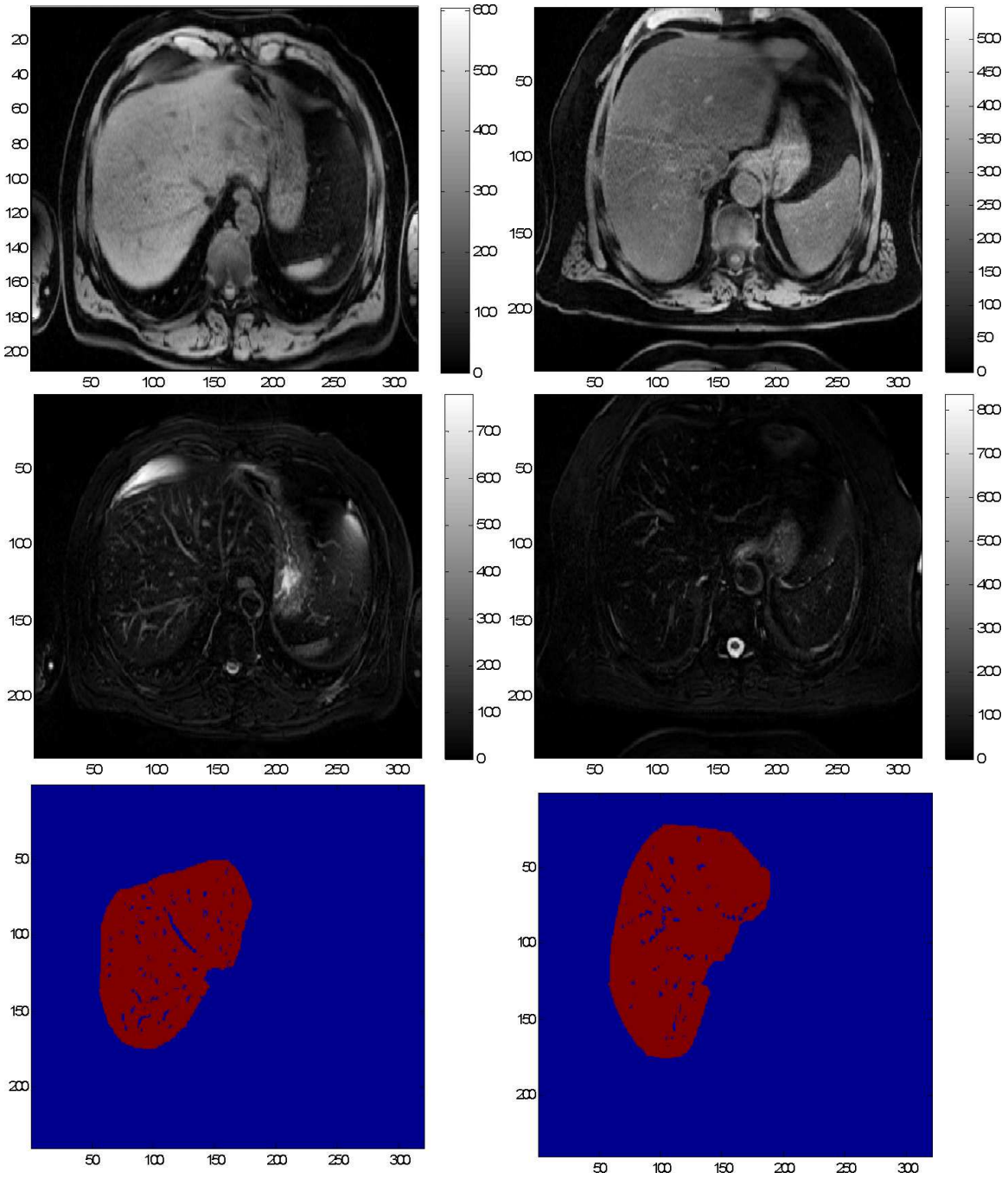


Figure 6.7: Vessel detection from T_2w images. The two columns correspond to RIAL05 and RIAL18. The first row is the T_1w image, the second row is the corresponding T_2w image, while the final row is the detected vascular and duct structures. Note the difference in vascular representation on the two different weighted images in favour of the T_2w. Note also the definite vessel structure around the x and y axes 100 and 100 on both images and their representation on the vessel detected masks in the final row.

The above algorithm was used to derive the liver masks in the following, unless otherwise indicated in the text or in the caption of the corresponding figure.

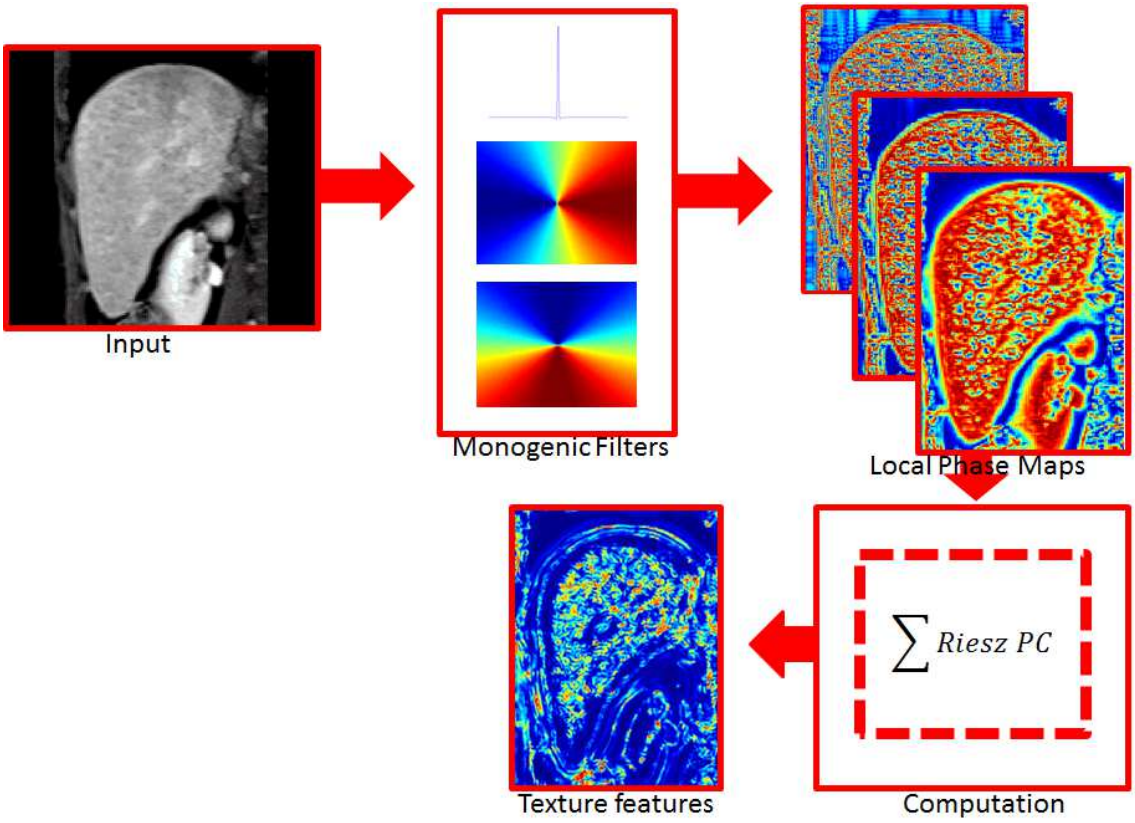


Figure 6.8: Multi-scale phase based texture analysis.

6.4 Local phase-based texture quantification with fractal dimension

We used the following steps to derive the texture score for each patient, as is illustrated in Figure 6.8:

1. Calculate the local phase at multiple scales using either the uncorrected or corrected intensity images;

2. Calculate the Riesz–weighted phase congruency that captures more of the internal texture of the liver than the boundaries;

3. Manually segment three representative slices of each patient; we choose to do this manually to avoid discussion over the possible effects of faulty classification results;

4. Derive the fractal dimension and compute the average over the three slices.

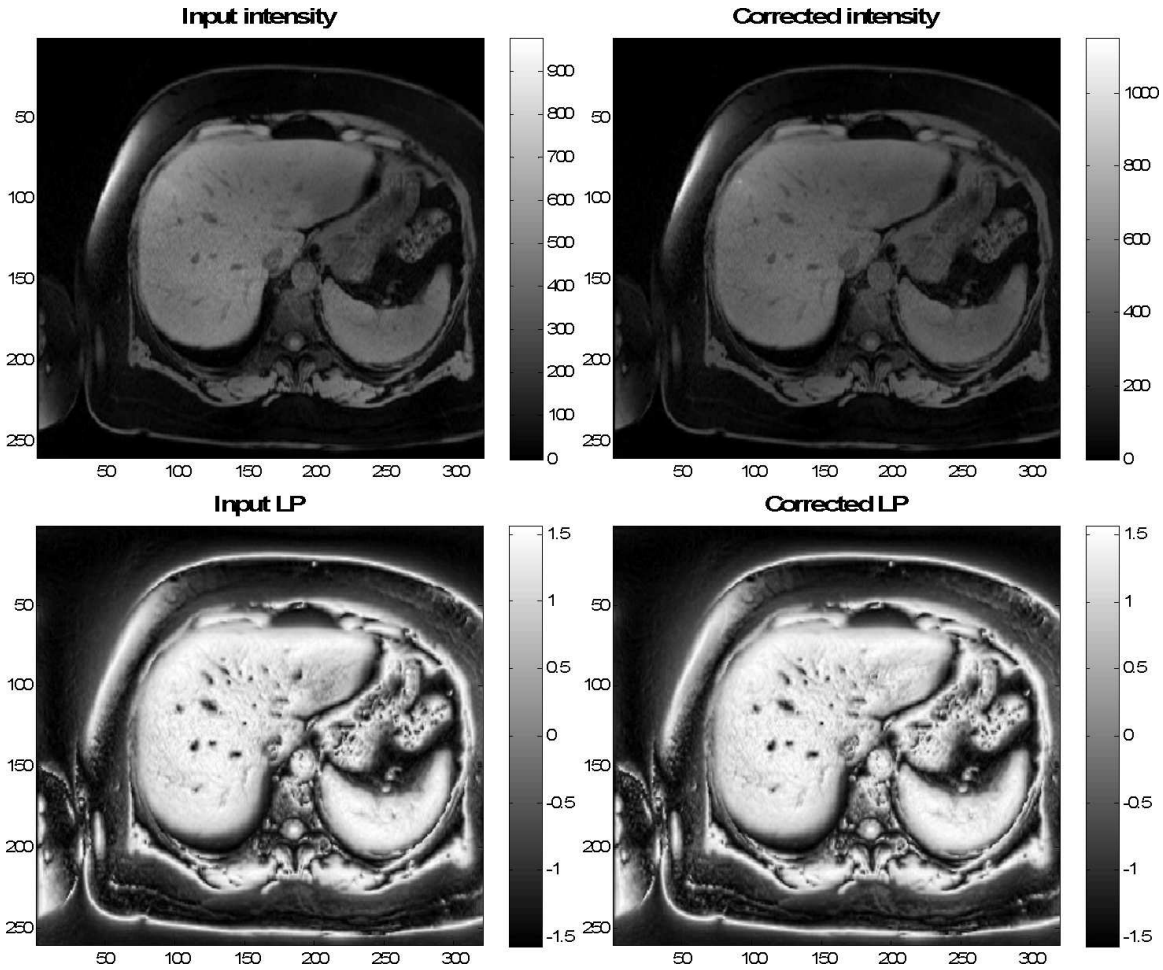


Figure 6.9: Comparison of the input and bias field corrected intensity (TOP) and the associated local phase maps (BOTTOM). By examination of the areas pointed at earlier in this chapter the regions that were corrected become obvious.

We have processed (15)19 axial liver Patient images in this way. We selected three slices from each patient. The criterion for selection was to maximize the liver tissue

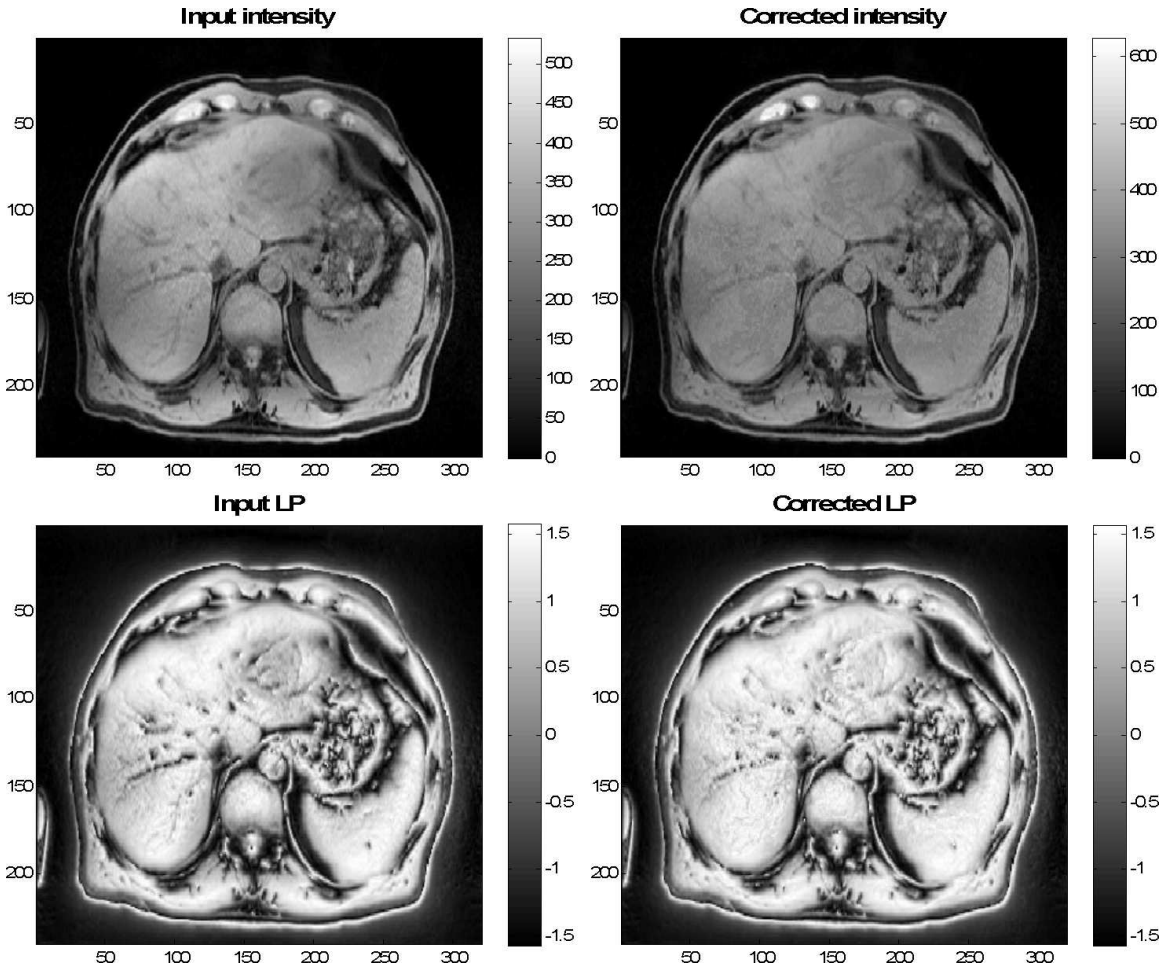


Figure 6.10: Bias field correction on Patient RIAL 10: The top two figures show the input intensity and the bias corrected intensity; the two images on the second row show the associated local phase maps. Note that the effect of bias field correction is more obvious on the phase maps.

and minimize other artifacts, for example motion which appears as blur. Nevertheless, these regions would be detected by our phase based measurements as having structural changes.

The intensity changes were corrected using the algorithm described in the previous section. Examples of corrected intensity and phase values are presented in Figure 6.9 and 6.10. The changes, that is to say the correction for the low intensity/phase patches, appear most clearly on the phase maps both on the top and on the left hand side of the axial slice.

The Riesz based phase congruency texture measures are then derived.

Finally, we derived the fractal dimension (FD) measures as a way to quantify the texture changes highlighted by the Riesz weighted phase texture congruency model. In Figure 6.11 we show the derived FD values for all 15 patients in segregating the clinically important three patient groups. All samples were bias-field corrected. Qualitatively the box-plots show the Riesz-weighted PC slightly outperforms the Kovesi type PC.

To estimate fractal dimension we used the well known box counting algorithm with a sliding window over six scales along the columns of the 2D Riesz PC texture measure. The particular non-linear filter used at different scales is defined as:

$$\mathrm{Ffilt}(r) = \left\lfloor \frac{\max(x) - \min(x)}{r} \right\rfloor \tag{6.11}$$

where $r = \overline{2,7}$ is the range of window sizes and x denotes the actual Riesz norm texture measure. We take the logarithm for each scale after filtering as:

$$B(r) = \log\left(\mathrm{Ffilt}(r)\frac{7^2}{r^2}\right) \tag{6.12}$$

The slope that defines the fractal dimension is then calculated using linear regression as:

$$FD = \frac{N_{xx}}{N_{xy}} \tag{6.13}$$

with

$$N_{xx} = i \star i - \frac{1}{6}\sum i^2 \tag{6.14}$$

$$N_{xy} = i \star B - \frac{1}{6}\sum i \star \sum B \tag{6.15}$$

where B is the collection of the filtered texture measures on different scales and $i \leftarrow \log(i)$, where i denotes scale. The FD values presented here are the average values derived from the three slices for each patient. The fractal lacunarity with the above notation was defined as:

$$FD_{lacunarity} = \sum FD^2 \cdot \frac{N}{(\sum FD)^2} - 1 \tag{6.16}$$

where N is the total number of pixels in the region of interest.

The fractal measure aims to quantify the local complexity in the sense that it measures to what extent the underlying signal has a nesting structure over multiple scales. These numbers may be fractional numbers. In general, structures that 'live' in a 2D space, a surface, are characterized by fractal values that are in the interval $[2-\epsilon, 2+\epsilon]$, where $\epsilon \in [0, 1]$ and it depends how self packed in a local neighbourhood they are.

There are three patient data-sets that are classified as having an associated FD value greater than two. These correspond to patients RIAL 5, 18 and 22. By visual examination of the actual liver intensity and phase, images RIAL 18 and 22 indeed have significant textural changes on both intensity and phase maps. However patient RIAL 5 mostly has an increased effect of blurring due to motion. The experiments performed within this framework show that the FD values, when calculated on the Riesz weighted PC, are a valuable source of descriptors (Figure 6.11), the median correlation is $r = 0.9667$ with a significance level $p < 0.0005$, however calculating the correlation with the seven scale Ishak scores shows very low correlation. As opposed to this the FD values based on the conventional energy weighted PC show very poor non significant correlations in both scenarios.

In addition to the FD values, the lacunarity of the structure was also computed. This measure quantifies the 'gapness' of the 2D surface of the liver in the derived

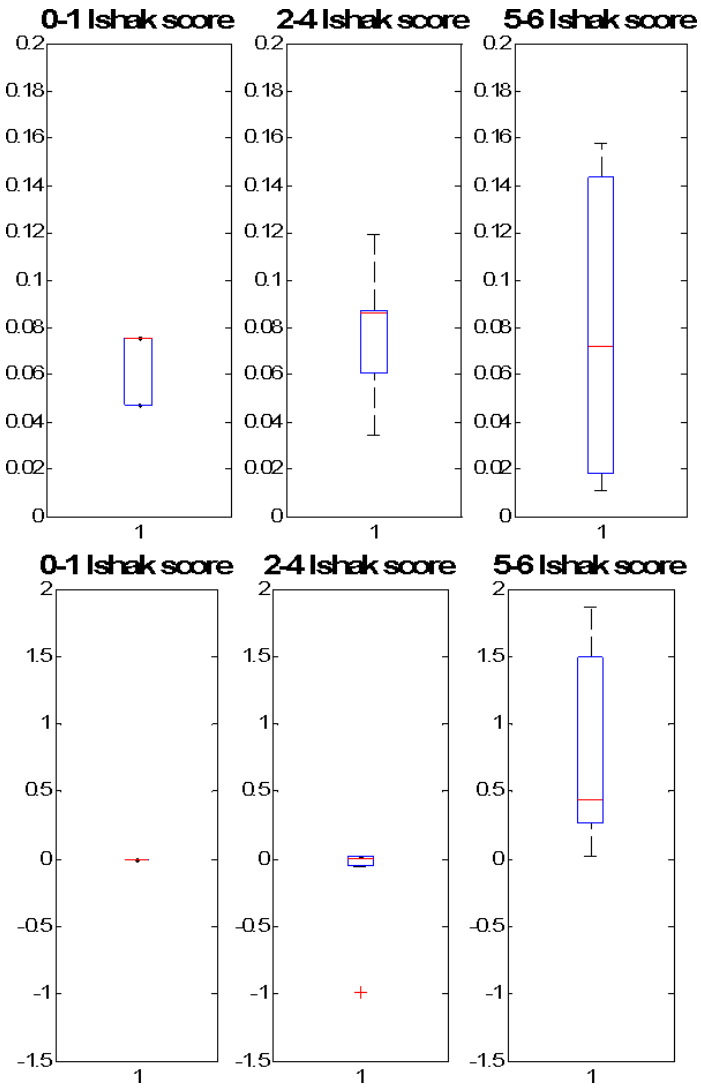


Figure 6.11: Fractal dimension (FD) statistics for the early, mild and severe disease. The boxplots show the median in red, the 25th and 75th percentile and the outliers with red crosses. The top line was derived based on Kovesi's phase congruency (PC), while the second row was derived based on the Riesz weighted PC. The correlation value for the median of the Kovesi type PC is $r = -0.3258$ which is not significant; the correlation between the derived scores and the full seven scale Ishak scores is $r = -0.3898$, which is also not significant. However, the Riesz weighted measure based median correlation is $r = 0.9667$ with $p < 0.0005$ significance level, and the correlation with the respective Ishak scores is as low as $r = 0.2596$, which is statistically not significant.

texture maps as is shown in Figure 6.12. Although the lacunarity values derived are based on the FD values, there is a non-linear relationship which shows up in the differences between the trends derived for the two measures.

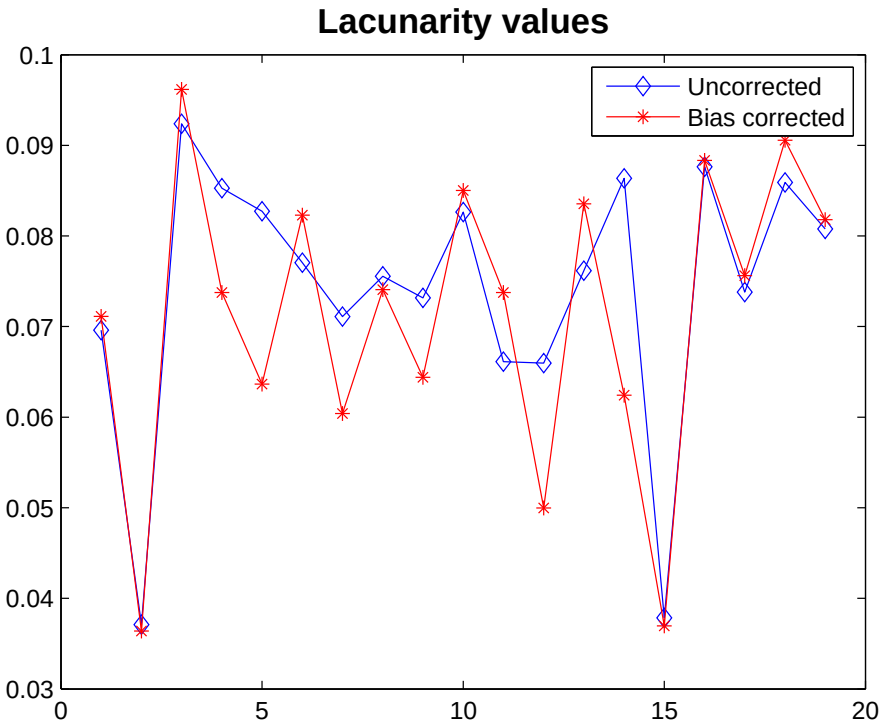


Figure 6.12: Fractal lacunarity values for the 19 patients. The blue denotes the lacunarity values calculated on the phase congruency texture measure without bias field correction of the input intensity image. The red data points are the results calculated with the exact same methodology, but with bias field correction of the input intensity values. No vessels were excluded during the derivation of these scores.

We have presented texture scores that we hypothesized to be surrogates for the state of the liver fibrosis using both average fractal dimension and lacunarity measurements. These are displayed in the above figures. We concluded that the Riesz weighted PC based FD values are valuable for tracking liver disease progression and so, as a next step, we evaluated the success of the methodology in separation of the pericellular and non-pericellular fibrosis.

6.5 Separation between the pericellular and non-pericellular fibrosis

Reasonably good separation between the pericellular and non-pericellular fibrosis for the respective 1, 2 and 6 Ishak scores may be obtained using the average fractal dimension as is shown in Figure 6.13. The separation is not perfect but it is nevertheless remarkable given that there has been no previous method to perform this automatically. The red line shows the results for the bias field corrected phase measurements while the blue line shows the results for the non corrected cases. The red arrows point to the cases with pericellular fibrosis and the blue arrows show the non-pericellular fibrosis. Except for Ishak score two there is clear separation between the two groups. For more details please see the caption of the figure.

The hypothetical threshold line was drawn by manual adjustment and currently we do not have a theoretical underpinning for it.

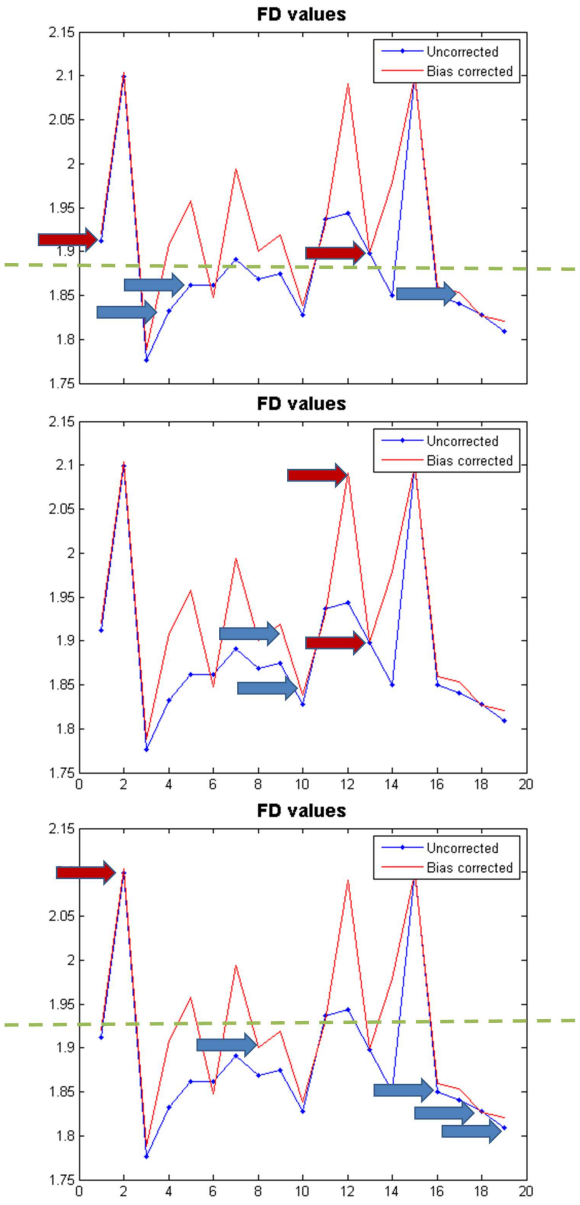


Figure 6.13: Separation of the pericellular and non-pericellular fibrosis from the average fractal dimension. The red arrows point to the cases with pericellular fibrosis, while the blue arrows are the ones with the non-pericellular fibrosis with the same Ishak score. From top to bottom the Ishak scores are as follows 1, 2 and 6. The x axis contains the patients in ascending order (as before).

6.6 Experimentation with closest distance entropy values

Entropy measurements come in many forms, either as global or local measurements, with variations on the map on which they are estimated. Since we are interested in the underlying texture we experimented with the entropy values on local signed distances motivated by the concept of the local binary patterns. The signed differences were calculated between neighbouring features and the order preserved (using an eight neighbourhood system).

The steps performed are as follows:

1. Calculate the difference between central pixels and their surroundings with an eight neighbourhood system. The question that arises is how to handle these eight signed differences. It is tempting to summarize them into one measure. However, making a sum of square differences or applying a similar measure is not adequate because it loses the order that is so important for the textures. Therefore all eight differences have been kept in the respective order they appear.

2. The eight neighbouring differences form different maps that are then processed. Each neighbourhood map has been used to calculate the NP windows pdfs and then the entropy is estimated from these pdfs. This is not necessarily the best measure, but is at least justifiable. This means that each liver is now associated with eight entropy values calculated on the signed differences from the eight neighbourhood systems.

3. Finally, we separate and rank them based on the eight length feature vector. The best way of achieving this remains a question, we used a statistical approach to do this.

Table 6.1: Entropy values for the 15 different patients, where the columns denote the different neighbourhood maps. For explanation on computation see text.

Patient iD	E1	E2	E3	E4	E5	E6	E7	E8
RIAL13	3.9541	4.2119	3.9933	4.1126	3.9532	4.2080	3.9968	4.1192
RIAL10	3.8309	4.1299	3.9168	4.0437	3.8315	4.1328	3.9209	4.0418
RIAL15	3.7903	4.0642	3.8623	4.0700	3.7945	4.0758	3.8706	4.0724
RIAL16	3.8763	4.1297	3.9514	4.1355	3.8831	4.1381	3.9537	4.1293
RIAL18	4.222	4.3919	4.2157	4.3831	4.2204	4.3889	4.2140	4.3817
RIAL21	3.9098	4.1402	3.9099	4.0812	3.9080	4.1296	3.9047	4.0747
RIAL11	3.8667	4.1581	3.9020	4.0699	3.8757	4.1701	3.9096	4.0653
RIAL17	3.9624	4.1753	3.9746	4.1438	3.9691	4.1818	3.9759	4.1374
RIAL07	3.6550	3.9175	3.6374	3.8845	3.6576	3.9274	3.6474	3.8875
RIAL22	4.2121	4.4509	4.2617	4.3489	4.2147	4.4487	4.2623	4.3479
RIAL06	4.3780	4.4868	4.3146	4.4881	4.3784	4.4801	4.3109	4.4819
RIAL14	3.8638	4.0777	3.8325	4.0315	3.8707	4.0802	3.8327	4.0206
RIAL23	3.6608	3.8872	3.6693	3.8190	3.6669	3.8957	3.6774	3.8180
RIAL28	3.6977	3.9035	3.7127	3.8917	3.7074	3.9259	3.7350	3.9039
RIAL30	3.4853	3.7593	3.5287	3.7172	3.4902	3.7571	3.5234	3.6990

The actual numerical values for each patient are displayed in the Table 6.1. There are 15 patients present instead of the 19 because these were the ones for which T_2w images for the vessel correction were available.

We used the box-plot measure to visualize the results, as is shown in Figure 6.14, which gives the assembled statistical information for the phase based texture measure based both on the energy weighted PC and Riesz weighted PC. The red line shows the median, the upper and lower horizontal line is the 25th and 75th percentile, the whiskers extend to the most extreme data and the red crosses (if any) are the outliers detected. The correlation between the median and group order was derived to be $r = 0.8359$ with a significance level of $p < 0.0005$ and a low correlation of $r = 0.3238$ between the actual texture measures and the Ishak scores (Figure 6.14). Similar good results were obtained with the energy weighted PC where the median based correlation is $r = 0.9197$ with a significance level of $p < 0.0005$ and a low correlation with the separate seven scale Ishak scores.

Slightly better results were obtained with the clumpedness measurement (Clark

and Evans, 1954) that takes into account spatial distribution. We used a modified version of the original formulation as detailed in the next section.

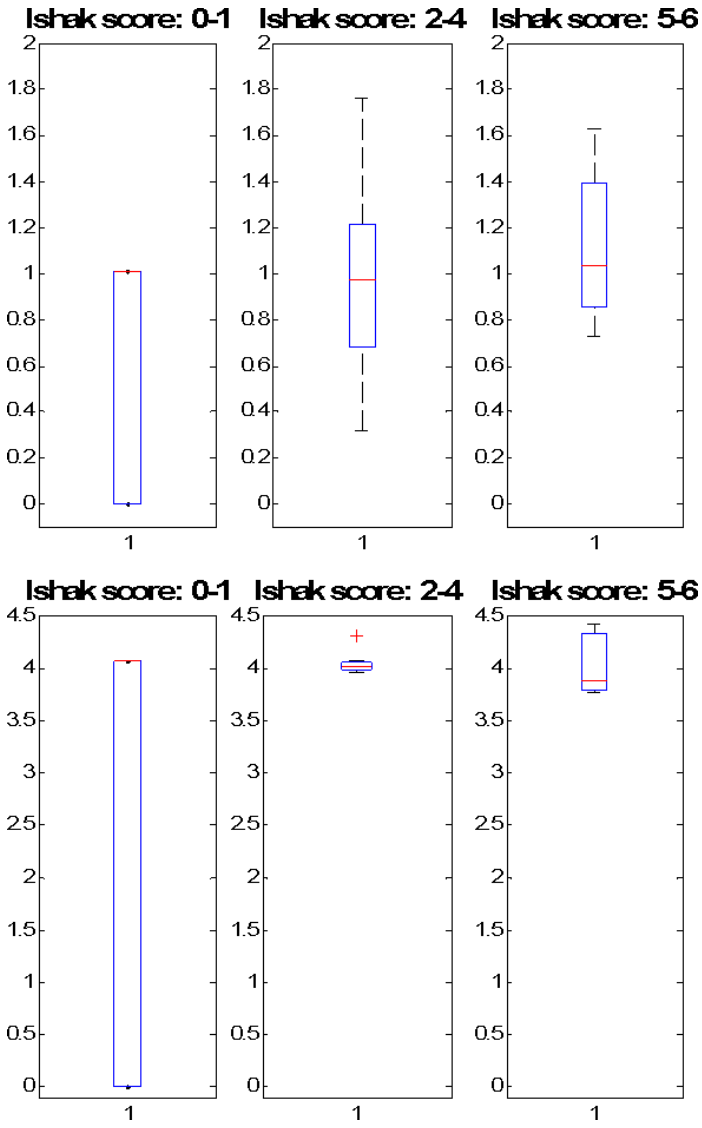


Figure 6.14: Statistics calculated on the closest distance entropy maps. The top row was derived based on the Kovesi's energy weighted phase congruency (PC), while the second row is based on the Riesz weighted PC. The boxplots show the median in red, the 25th and 75th percentile and the outliers with red crosses. Based on the energy weighted PC the correlation for the median of the three group is $r = 0.9197$ with a significance level of $p < 0.0005$, and the correlation with the respective Ishak scores is $r = 0.2617$, which is statistically not significant. Using the Riesz weighted PC the correlation for the median is $r = 0.8359$ with a significance level of $p < 0.0005$, and correlation with the Ishak scores is $r = 0.3238$, which is statistically not significant.

6.7 Clumpedness measurement

Fibrosis initiates in spatially separate clusters, so measurements that take into account spatial locality are to be preferred. One way to assess the texture of the underlying Riesz and energy weighted liver images is to use the clumpedness spatial statistical measure originally proposed in (Clark and Evans, 1954). Direct application of the original measure is hard because there is no obvious way to measure certain quantities in a 2D image that were initially derived for animal populations. Therefore it is calculated as:

$$\text{clutter} = \frac{\frac{1}{n}\sum \text{closestdist}}{\sqrt{\text{density}}} \tag{6.17}$$

where n denotes the total number of closest distance measurements, and density is not well defined. In case of a random distribution (which is not the subject of the current discussion) it becomes the expected value of the inter-species length. Instead of density, we have used the area of the liver to infer a measure of density. With this measure we obtained the results shown in Figure 6.15.

At this stage we do not expect a perfect trend with the Ishak scoring system because these themselves were made based on hepatologist opinion instead of a real ground truth.

Figure 6.15 shows the box-plot analysis for the phase based texture analysis, both the energy and the Riesz weighted PC. The box-plot shows the 25th and the 75th percentile (two horizontal limits), the median in red and the whiskers show the data points that extend far beyond the expected but they cannot as yet been classified as being outliers. Outliers are denoted by red crosses. To examine the value of the measure we have derived statistics for the correlation of the median with the selected three groups and the correlation between the measures with the individual seven scale Ishak scores. The Riesz weighted PC based measure has a correlation of $r = 0.9580$

with a significance level of $p < 0.0005$ with respect the group median and a correlation of $r = 0.3672$ with respect to the Ishak scores. Based on the energy weighted PC the correlations for the median become $r = 0.8189$ with a significance level of $p < 0.0005$ and a correlation with the Ishak scores of $r = 0.3458$. Finally, we have assessed the well known Law's measures in the following closing section.

6.8 Laws texture features

We also evaluated the Laws texture measures on the available liver images. These rely on five 1D kernels and the 2D kernels built from them. By an appropriate combination of kernel responses, the measures can be made rotationally invariant for almost all features. The 1D kernels are as follows:

$$\begin{aligned} L_5 &= [1 \quad 4 \quad 6 \quad 4 \quad 1] \\ E_5 &= [-1 \quad -2 \quad 0 \quad 2 \quad 1] \\ S_5 &= [-1 \quad 0 \quad 2 \quad 0 \quad -1] \\ W_5 &= [-1 \quad 2 \quad 0 \quad -2 \quad 1] \\ R_5 &= [1 \quad -4 \quad 6 \quad -4 \quad 1] \end{aligned}$$

which correspond to level, edge, spot, wave and ripple. The advantage of these measures is that extraction of texture features is possible with the application of concomitant fast convolution filters, without demanding processing and memory usage. The computation proceeds as follows: the kernels are applied to the images and then the local texture energies are calculated in a neighbourhood of 7×7 on the filtered images. These are called texture energy maps, in short TEMs, and they are normalized by a nonzero sum kernel response to lead to the feature map. Corresponding normalized TEMs are summed to yield rotationally invariant feature maps. Texture

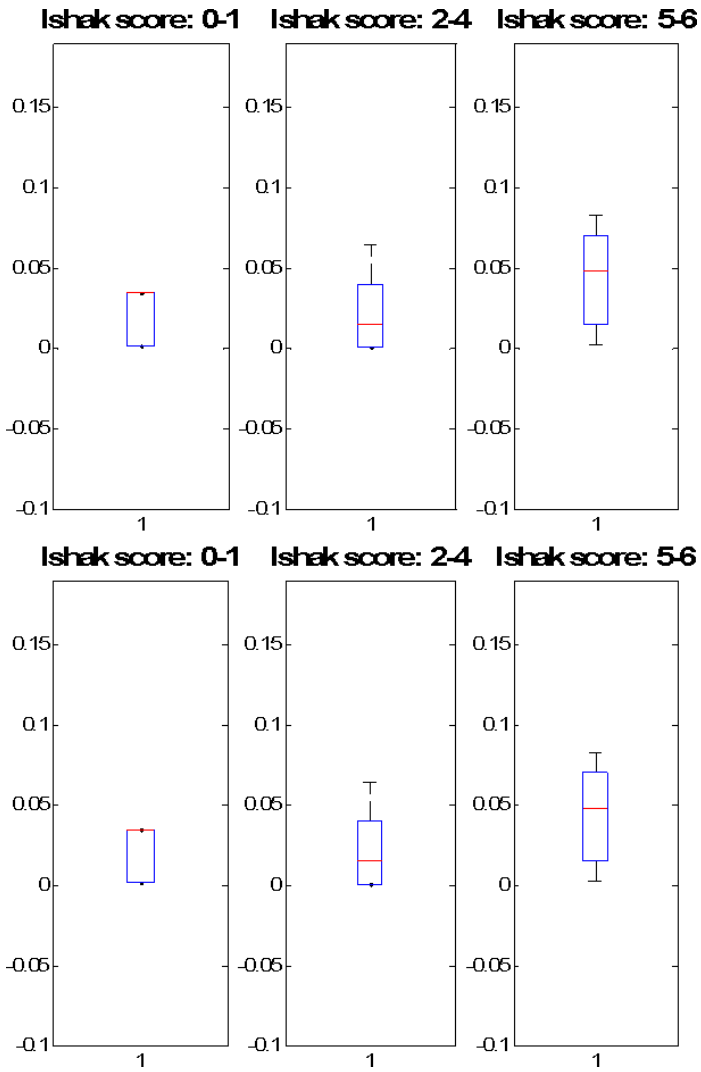


Figure 6.15: Statistics based on the clumpedness measurement. The top row was derived based on the Kovesi's phase congruency (PC) and the second row is based on the proposed Riesz weighted PC. The boxplots show the median in red, the 25th and 75th percentile and the outliers with red crosses. Using the energy weighted PC the correlation based on the median is $r = 0.8189$ with a significance level of $p < 0.0005$, the correlation of the derived measures with the Ishak score $r = 0.3458$, which is statistically not significant. Based on the Riesz weighted PC the correlation for the median $r = 0.9580$ with a significance level $p < 0.0005$, and the correlation with the Ishak scores is $r = 0.3672$, which is statistically not significant.

classification based on normalised TEMs are the standard way today to calculate classification of features based on image intensities. However, in general, they tend to lead to limited information in case of the intensity based medical images. Instead we used the Laws measures based on derived energy and Riesz weighted PC maps.

Figure 6.16 shows the statistical analysis of the grouped data that correspond to no disease (Ishak score 0-1), to mild disease (Ishak score 2-4) and to severe fibrotic disease (Ishak score 5-6). We made the particular choice between the grouping to assess clinically meaningful separation between the group, because a stage 4 liver disease is still 'treatable'. We calculated the texture feature summaries in the form of average values. We compare the results between the Riesz weighted and energy weighted PC. The Riesz based texture measure provides significant correlation both respect the median of the three clinically relevant disease groups and the actual Ishak scores. The median based correlation is $r = 0.8778$ with a significance level $p < 0.005$ and the correlation of the measure with the respective Ishak scores is $r = 0.5790$ with a significance level of $p < 0.01$. This is the best result out of the measures we have evaluated in this chapter. The energy based texture map provides a median based correlation of $r = 0.42$ only with a significance level of $p < 0.05$ and a correlation with the Ishak scores of $r = 0.1823$. Interpretation of the boxplots is as follows: the red horizontal line shows the mean value detected in the respective group, the edges correspond to the 25th and 75th percentile of the data, while the whiskers extend to the most extreme data points, which are however not considered as yet outliers.

Figure 6.17 shows an alternative representation of the trend but this time we used a different visualization to aid understanding. The three groups correspond to the exact same segregation as in case of the previous image: Ishak 0-1. Ishak 2-4 and Ishak 5-6 corresponding to no disease, mild and severe fibrosis. As expected there are overlaps between the groups and there are also outliers unfortunately not detected with the box-plot method, but the main trend is clearly observable.

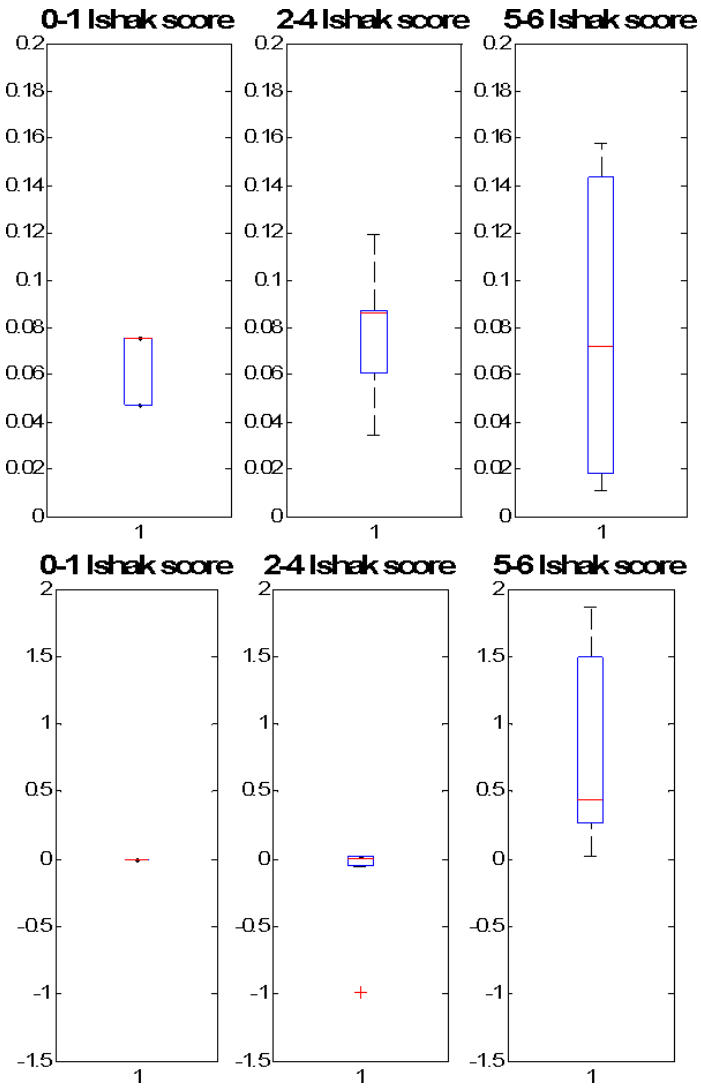


Figure 6.16: Statistics based on Laws texture measures. The first row was derived based on the Kovesi type PC, while the second row is based on the Riesz weighted PC. The boxplots show the median in red, the 25th and 75th percentile and the outliers with red crosses. Based on the energy weighted PC the correlation of the median with the respective three groups is $r = 0.42$ with a significance level of $p < 0.05$, and the correlation between the derived measures with the Ishak scores is $r = 0.1823$ which is statistically not significant. Using the Riesz weighted PC the correlation between the group median and the respective group order is $r = 0.8778$ with a significance level $p < 0.0005$, and the correlation of the derived scores with the Ishak values is $r = 0.579$ with a statistical significance $p < 0.01$.

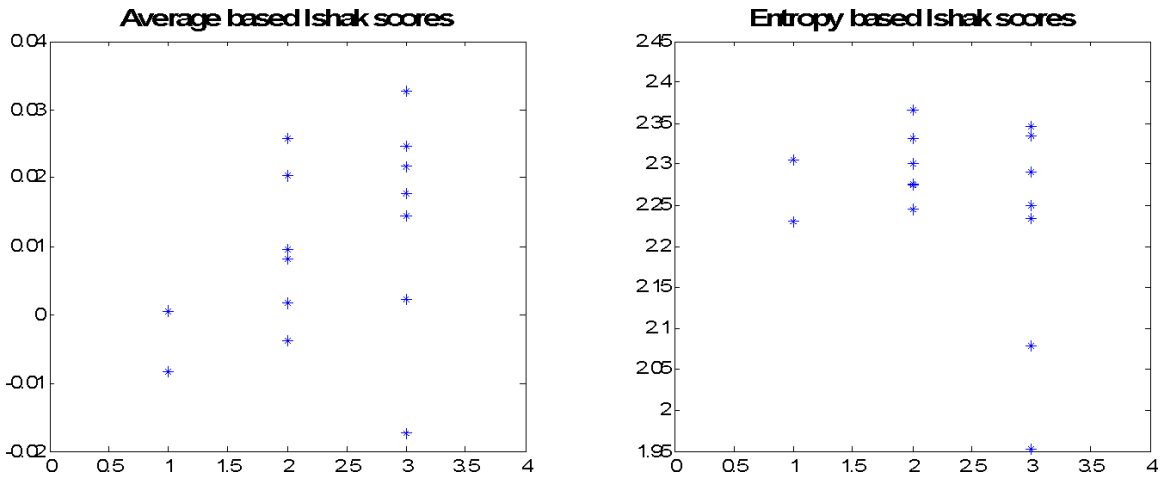


Figure 6.17: Laws texture measure in detecting the Ishak trends using the individual average and entropy based measurements. This is a different way to visualize the trends and which reflect the statistically found ones.

6.9 Kovesi versus Riesz-weighted PC

Finally, we have assessed the characterization power of the Kovesi and Riesz weighted PC without the application of a secondary spatial filter as we have presented earlier in this chapter. The feature vectors were summarized in 256 length non-parametric probability density functions. Then linear discriminant analysis (LDA) was performed to separate those patients who would not benefit from a potential treatment, i.e. Ishak 5-6 from the rest of the group.

The quantitative evaluation is summarized in Table 6.9 in terms of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). These were implemented in Matlab 2013a and included 15 patients without bias field correction, but with vessel removal.

The results suggest that the proposed method performs as well or better than the existing Kovesi type PC.

Table 6.2: Linear Discriminant Analysis: Ishak 0-4 and Ishak 5-6 patient groups based on texture PDFs.

Measure/Method	Sensitivity	Specificity	PPV	NPV
Kovesi PC	0.8571	1.0	0.8571	0.9231
Riesz PC	1.0	0.9286	1.0	1.0

6.10 Conclusion

Several conclusions may be drawn from the experimental results shown in this Chapter. First, all aspirations that we made in the beginning of the chapter are possible in some form but not with the same measure.

Fractal measures and even fractal likeness very rarely holds. However, it nevertheless gives a measure of complexity and has proved to be a valuable source of descriptor both for the trend during disease progression and the separation between pericellular and non-pericellular disease - that cannot at current be done by any other methodology.

The measures presented in this chapter additionally include the local entropy based on signed differences, the clumpedness measurement, and the Laws texture measures. All these measures were able to detect the trend of the median of the three disease groups. Meaningful Ishak score trends are best found by the Laws texture measures based on the Riesz weighted PC measurements. This was the only measure that shows significant correlation both for the median of the three groups and with the Ishak scores.

Overall, all the derived statistical measures include as comparison the calculations based on the energy weighted PC maps. This reflects our assertion that the Riesz weighted PC systematically outperforms the energy weighted PC especially for the Laws texture measures.

Finally, the pure PC value probability density functions show, that the Riesz-weighted PC performs at least as well as the Kovesi PC, or better, in differentiating

the patient group that would not benefit from further treatment.

Note: our measurements were based on a limited number of available low-quality patient MRI images, therefore the extrapolation of our results are subject to potential sampling errors.

Chapter 7

Conclusion and future work

In this final chapter, we first summarise what we consider that we have accomplished in this thesis, then we outline possible research directions for future work that arise from the research presented in this thesis.

7.1 Summary

We set out to differentiate between different stages of liver disease from suboptimal T_1w liver images. To do this, we proceeded on several fronts. First, we developed meaningful texture measure energies. Second, we assessed current statistical descriptors of these energies; and finally, we investigated the confidence of current state-of-the-art local orientation estimators. Having these in place, we applied the resulting improved methods to real liver images to test their differential power as compared to Kovesi's energy weighted PC.

We have evaluated methods to improve the current continuous intrinsic dimensionality formulation. To do this, we first investigated local orientation estimators. This evaluation suggested that although there exist more precise estimation methodologies than the monogenic signal, the latter is predictable, so it also desirable in many applications. A precise uncertainty model based on the above uncertainties has

been derived for the ciD model in its entirety.

The most burning question addressed in this thesis is feature/texture energy. This was designed to be intensity and contrast change independent, therefore local phase based. Within this framework we developed the Riesz weighted phase congruency. This measure gives very promising results in detecting internal textural information, but it is hopeless in the delineation of organ boundaries. This measure becomes the basis for our liver disease differentiation.

Finally, we have shown that texture measures calculated on the derived texture energy based on local phase has many advantages relative to energy weighted PC. The measurements based on fractals, entropy, clumpedness and Laws energies showed that it can differentiate the main trends in Ishak scores even from these suboptimal liver MRI scans. What is more, the fractal measure was able to differentiate to a large extent between pericellular and non-pericellular fibrosis. This latter is an important contribution because no other algorithm currently exists that can do this.

7.2 Future work

Based on the results presented in this thesis and observed during the development of the methodologies we summarize our perspective on future research directions.

The proposed Riesz phase congruency model is a good case to demonstrate the idea on how features can be extracted using a phase based method, however it currently has several drawbacks: local phase values are not in general defined for non-i1D features and their uncertainty increases with the isophote curvature of any i1D line segment. This holds true for the different forms of the monogenic signal, such as the conformal monogenic signal; since this problem is inherently embedded in the signal model itself the question is whether using either the two Riesz filters or the conformal embedding it is possible to surpass this obstacle. The two challenges that we are

faced with is to first arrive to an agreement (definition) of phase values for non-i1D features and then to devise a signal model to capture this property.

Textures inherently exist on multiple scales, therefore the use of the derived features to define a generic texture measure is essential. To-date we have used a selected characteristic scale and several scales around this scale. Also, we have used fractal measure, the Laws texture filters, entropy and clumpedness with good results on the liver samples. To address this question there are several ways forward: either to use an isotropic scale-space (for example diffusion led Gaussian based (Lindeberg, 1996)), to use an anisotropic Perona-Malik (Perona and Malik, 1990) type scale-space or to use a specific scale-space for example by sampling the space with a circle or ellipsoid (Kadir and Brady, 2001). There are a third class of multi-scale possibilities, which includes something along the lines of an inclusion tree (Hong, 2004), with the anisotropic scale-space (Kadir, 2002) or more recent is the texture measure in (Hong et al., 2011) based on local entropies on multiple scales with NP windows.

The evaluation of the texture measures may be made through different approaches. We have so far only evaluated the methods targeting the texture of the liver parenchyma in specific, as is presented in the thesis. To test the general applicability of the devised methods we plan in the future to test these using selected texture databases, such as the Brodatz or VisTex. In addition, we have not evaluated different distance measures to compare the derived measures. This would also allow to build an ROC curve to be able to compare the performance of the texture descriptors to those of the one presented in (Mellor et al., 2008) and in (Lazebnik et al., 2005).

Specific observations for the characterization and classification of liver samples

The bias field correction presented in the thesis is not optimal for the liver magnetic resonance images (MRI) we have. The first task would be to implement Steyner's method based on local polynomials (Styner et al., 2000) and to devise a method

that is not based on the piecewise smooth image model. The fact that we applied a piecewise smooth model greatly reduces the computational complexity, however it does not hold true for most of the abdomen despite its wide use. One observation made in this thesis is that local phase is severely influenced by the presence of the bias field. To our knowledge, bias field correction based on local phase values have not been explored in the literature. Since both the intensity and the local phase is affected by the bias field to a different degree, the particular question is how to link these two values to come-up with a joint detection. One particular possibility is the use of zero phase filters, which allows that a logarithmic connection between the intensity magnitude and the phase to be made. Another possibility is to use the exponential wavelet signal model based on octave band-pass filters proposed in (Anderson, 2007).

The segmentation of the liver from the 2D slices is currently performed manually and it is particularly challenging to do this automatically on these suboptimal images, whether the method is 2D or 3D based. There are two main tasks to be overcome: natural inhomogeneity of the liver parenchyma; the presence of blood vessels and bile ducts possibly their segmentation. Another obstacle is the non-continuous or disrupted liver edge treatment, this could be overcome with methods developed in the computer vision community with respect subjective contours. However, we do not expect this to be applicable straight away because the images we are faced with constitute a more challenging problem. This is because the boundary is ill defined to start with and its surrounding is often similar to that of the liver (for example, some parts of the rectum and the lower part of the heart may present similarly on the MR images).

To be able to extract the blood vessels and bile ducts it is best to use an MRI weighting that is sensitive to fluid. This is provided by the T_2w images. However, these images are not optimal for the texture quantification. Therefore the registration

of the T_2w and T_1w images is necessary. To do this in the best possible way we first aim to re-sample the voxels to an isotropic resolution, for example using the Riesz fusion developed by Wen Jun Chi in 2012. Once this is done an affine registration may be performed using for example the method developed by Daniel Rueckert and Julia Schnabel.

Novel ciD method and ciD based noise estimation

One of the first observations made in the implementation and testing of the continuous intrinsic dimensionality is that although the probabilities derived are meant to represent confidences, themselves are prone to inaccuracies, which in fact may be described by geometric structures. The thesis presented a more general uncertainty description based on Xavier Pennec's earlier proposed framework that applies to the derivation of uncertainties in a general framework. Instead here we propose that a specific geometric structure, namely a truncated cone may be used to quantify this as is presented in the figure in Chapter 5. The parametrization of this depends on two main measurements based on the origin variance and line variance. In the thesis it was pointed out that there is subtle connection between the double orientation and the histogram oriented gradients (HOG) (Dalal and Triggs, 2005), both of which have their advantages and disadvantages. The double angle representation allows the inclusion of spatial neighbourhood relations, while the HOG is better in capturing the energy associated with each orientation.

The different iD structures are differently represented in various transform spaces. In specific, there are significant regular representations between the Fourier and spatial domain. For example, while in the spatial domain the i0D spans the whole space and it is fully predictable at each position, it is coded by a delta peak multiplied by some constant in the Fourier domain. One particular exciting approach would be to use an iterative approach in which both the spatial domain and the Fourier domain descriptors are used with their constraints.

As noted in this thesis there are a series of shortcomings all of which needs further development: the definition of an accurate energy measure may be dependent on the imaging modality to which it is applied, at the same time so far we have only used the i1D local orientation, both of which needs further assessment. Although the ciD triangle is a remarkable advance in computer vision, and especially in low level feature detection, there are two aspects that need further consideration:

1. Neither our experiments, nor results presented in (Kalkan et al., 2007) or (Felsberg et al., 2009) show that there is a monotonic change along the vertices of the triangle. This is presumably due to the fact that the method relies on fundamentally unsolved computer vision problems with nonlinear changes.

2. Curiously it is suggested that the textures occupy a relatively small portion of the triangle along the ci1D-ci0D transition. This may well be due to incorrect local orientation estimation or scale selection, but not necessarily due to real textures.

During experiments not presented in this thesis it was observed that the ciD coordinate movement as a result of increasing Gaussian noise does not follow a linear trend within the triangle itself, instead, it is described by a non-perfect projectile motion. This observation was made by adding increasing amount of Gaussian noise to ideal samples representing the three vertices of the ciD triangle. The nonlinear motion is expected to be a complex result of the interplay between the three ciD coordinates as a result of added noise.

The ciD of the same feature changes on different scale, to derive a method that faithfully represents the textures it is desired to devise a multi-scale method that derives these based on a set of characteristic scales, methods listed above for the phase based evaluation could be used as a first reference point.

Publication list

Journal papers

1. R. Ali, T. Szilágyi, M. Gooding, M. Christlieb, M. Brady,"On the use of low pass filters for image processing with inverse Laplacian models", Journal for Mathematical Imaging and Vision, 2011.

2. R. Ali, M. Gooding, T. Szilágyi, B. Vojnovic, M. Christlieb, M. Brady,"Automatic segmentation of adherent biological cell boundaries and nuclei from brightfield microscopy images", Machine Vision and Applications, 2011.

3. R. Ali, G.-D. Cigdem, T. Szilágyi, B. Durkee and E. Graves, Semi-automatic segmentation of subcutaneous tumours from micro-computed tomography images." PMB. IN Press.

Peer-reviewed conference papers

1. T. Szilágyi, M. Brady, "Feature extraction from cancer images using local phase congruency: a reliable source of image descriptors", International Symposium on Biomedical Imaging, Boston, 28th June-1st July, 2009.

2. T. Szilágyi, M. Brady, T. Brunner, N. Joshi, "Local phase significance estimated with uncertainties to detect fibrotic regions from in vivo pancreatic cancer images", Medical Image Understanding and Analysis, London, 14th-15th July, 2009.

3. T. Szilágyi, M. Verhoek, A. Noble, "Detecting early response to therapy in liver cancer treatment: 3D metastases segmentation using graph-cuts with a modified prior", Medical Image Understanding and Analysis, London, 14th-15th July, 2009.

4. T. Szilágyi, A.K. Harvey, M. Brady, L. Cochlin, N. Joshi, "Quantitative analysis of tendon ECM damage using MRI", International Symposium on Biomedical Imaging, Rotterdam, Netherlands, 14-17 April, 2010 (POSTER).

5. R. Ali, T. Szilágyi, K. Hussien, M. Christlieb, B. Vojnovic, M. Brady, "Automatic nucleus segmentation of adherent cells from brightfield microscopy images", British Machine Vision Association and Society for Pattern Recognition Symposium, London, 21 April, 2010.

6. B. Cornelissen, S. Able, T. Szilágyi, K. Shah, M. Brady, K. Vallis, "DNA double strand break imaging with anti γH2AX-TAT antibodies does not alter DNA repair kinetics", AACR Annual Meeting, Chicago, 31 March 4 April, 2012 (POSTER).

7. R. Ali, T. Szilágyi, E. Graves, "Semi-automatic phase-based segmentation of subcutaneous tumours from microcomputed tomography images", International Symposium on Biomedical Imaging, Barcelona, 2-5 May, 2012.

Workshop Talks

1. M. Brady and T. Szilágyi, "The monogenic signal, Section: Robust Feature Descriptors", Local Structure Workshop, Manchester, 5 May 2009.

2. T. Brunner and T. Szilágyi, "Ultrasound imaging of preclinical tumours focusing on the microenvironment (vascularity and stroma)".Invited Talk, British Medical Ultrasound Society Pre-clinical Day, London, 20 November, 2009.

Bibliography

Anatomy, M. and System, B. (2013). Liver ductular system. `http://imageradiology.blogspot.com`. [Accessed on 19-09-2013].

Anderson, R. (2007). *Phase based object matching using complex wavelets*. PhD thesis, University of Cambridge, UK.

Asselah, T., Bieche, I., Sabbagh, A., Bedossa, P., Moreau, R., Valla, D., Vidaud, M., and Marcellin, P. (2009). Gene expression and hepatitis C virus infection. *GUT: Recent advances in basic science*, 58:846–858.

Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press.

Bloch, F. (1946). Nuclear induction. *Physical Review*, 70:460–474.

BLT (2008). British liver trust: Facts about liver disease. `http://79.170.44.126/britishlivertrust.org.uk/home-2/media-centre/facts-about-liver-disease/`. [Accessed on 19-09-2013].

Bracewell, R. (2000). *The Fourier Transform and its applications*. McGrow-Hill.

Bruske, J. and Sommer, G. (1998). Intrinsic dimensionality estimation with optimally topology preserving maps. *IEEE TPAMI*, 20:572–575.

Canny, J. (1986). A computational approach to edge detection. *IEEE TPAMI*, 8(6):679–698.

Caselles, V., Kimmel, R., and Sapiros, G. (1997). Geodesic active contours. *IJCV*, 22(1):61–79.

Centre, M. M. I. (2011). Could mushrooms help with kidney and liver disease? `http://medicinalmushroominfo.com/could-mushrooms-help-with-kidney-and-liver-disease/`. [Accessed on 19-09-2013].

Clark, P. J. and Evans, F. C. (1954). Distance to nearest neighbour as a measure of spatial relationships in population. *Ecology*, 35:445–453.

Cohen, M. M. S. (2006). Merck manuals, online. `http://www.merckmanuals.com/professional/index.html`. [Online; accessed October-2012].

Comaniciu, D. and Meer, P. (2002). Mean shift: a robust approach towards feature space analysis. *IEEE TPAMI*, 24:603–619.

Curtis, S. (1985). Reconstruction of multidimensional signals from zero-crossings. Technical report, Massachussetts Institute of Technology, Research Laboratory of Electronics, Report No. 509.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 886–893.

Dall'Armellina, E., Piechnik, S., Ferreira, V., Si, Q., Robson, M., Francis, J., Cuculi, F., Kharbanda, R., Banning, A., Choudhury, R., Karamitsos, T., and Neubauer, S. (2012). Cardiovascular magnetic resonance constrast T1-mapping allows assesment of severity of injury in acute myocardial infarction. *Journal of Cardiovascular Magnetic Resonance Imaging*, 14:15–27.

Day, C. (2011). Non-alcoholic fatty liver disease: a massive problem. *CME Hepatology*, 11:2:176–8.

Delakis, I. (2006). Spatial encoding and MRI sequences. [Teaching Notes for the Department of Bioengineering, Imperial College London].

Destrempes, F. and Cloutier, G. (2010). A critical review and uniformized representation of statistical distributions modelling the ultrasound echo envelope. *Ultrasound in Med. and Biol.*, 36:1037–51.

Felsberg, M. (2002). *Low-level image processing with the structure multivector*. PhD thesis, Christian-Albrechts-University, Kiel.

Felsberg, M., Kalkan, S., and Krüger, N. (2009). Continuous dimensionality characterization of image structures. *Image Vision Comput.*, pages 628–636.

Felsberg, M. and Köthe, U. (2005). Get: The connection between monogenic scale-space and Gaussian derivatives. In Kimmel, R., Sochen, N., and Weickert, J., editors, *Scale Space and PDE Mathods in Computer Vision*, LNCS, pages 192–203. Springer.

Felsberg, M. and Krüger, N. (2003). A probabilistic definition of intrinsic dimensionality for images. In *LNCS: DAGM*, pages 140–7.

Felsberg, M. and Sommer, M. (2002). The monogenic signal. *IEEE TSP*, 49:3136–3144.

FerriScan (2009). Whittington Hospital, NHS.

Field, D. and Brady, N. (1997). Visual sensitivity, blur and the sources of noise variability in the amplitude spectra of natural scenes. *Vision Research*, 37(23):3367–83.

Fierbinteanu-Braticevici, C. (2009). Non-biopsy methods to determine liver fibrosis. `http://www.medandlife.ro/medandlife346.html`. [Accessed on 19-09-2013].

Fitzgerald, W. (1991). Extension of Logan's theorem. Technical report, University of Camridge, UK.

Fleischmann, O., Wietzke, L., and Sommer, G. (2011). Image analysis by conformal embedding. *J. Math. Imaging Vis.*, 40:305–325.

Förstner, W. (1998). Image preprocessing for feature extraction in digital intensity, color and range images. In *Proceedings of the International Summer School on Data Analysis and the Statistical Foundations of Geomatics, Lecture Notes on Earth Sciences*. Springer.

Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Academic Press.

Gabor, D. (1946). Theory of communication. Part 1: The analysis of information. *Journal of the Electrical Engineers- PART III: Radio and Communiaction Engineering*, 93:429–441.

Gareau, P., Wymore, A., Cofer, G., Karlik, S. J., and Johnson, G. (2001). Imaging inflammation: Visualization of perivascular cuffing in EAE by Magnetic Resonance Microscopy. *Proc. Intl. Soc. Magn. Reson. Med.*, 9:1.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE PAMI*, 6:721–741.

Geusebroek, J.-M. (2005). The stochastic structure of images. In Kimmel, R., Sochen, N., and Weickert, J., editors, *Scale-space*. Springer-Verlag Berlin Heidelberg.

Gilles, S. (1998). *Robust Description and Matching of Images*. PhD thesis, University of Oxford.

Giorgio, V., Prono, F., Graziano, F., and Nobili, V. (2013). Pediatric non alcoholic fatty liver disease: old and new concepts on development, progression, metabolic insight and potential treatment targets. *BioMed Central Pediatrics*, 13:40:1–10.

Gosta, G. and Knutsson, H. (1995). *Signal processing for computer vision*. Springer.

Gould, T. (2013). MRI magnets: the major players. `http://science.howstuffworks.com/mri1.htm`. [Accessed on 19-09-2013].

Granlund, G. and Knutsson, H. (1995). *Signal Processing for computer vision*, chapter Orientation and velocity/ Representation and averaging. Kluwer Academic Publisher.

Guillemaud, R. and Brady, M. (1997). Estimating the bias field of MR images. *IEEE TMI*, 16:3:238–251.

Hayes, M. H. (1981). *Signal reconstruction from phase or magnitude*. PhD thesis, Massachussetts Institute of Technology.

Holmes, J. and Bydder, G. (2005). MR imaging with ultrashort TE (UTE) pulse sequences: Basic principles. *Radiography*, 11:163–174.

Hong, B.-W. (2004). *Image segmentation, shape and registration: application to mammography*. PhD thesis, University of Oxford.

Hong, B.-W., Ni, K., and Soatto, S. (2011). Entropy-scale profiles for texture segmentation. *SSVM*, pages 243–254.

Hurt, N. (1989). *Phase retrieval and zero-crossings: mathematical methods in image reconstruction*. Kluwer Academic Publisher.

Institute image, T. S. and video galery (2009). Gpu based level-set methods used for segmentation of liver vasculature in mouse thorax mri data. `http://www.sci.utah.edu/gallery2/v/cibc/mouseReg.tif.html`. [Accessed on 19-09-2013].

Iredale, J. (2007). Models of liver fibrosis: exploring the dynamic nature of inflammation and repair in a solid organ. *Journal of Clin. Investigation*, 117:539–548.

Ishak, K., Baptista, A., Bianchi, L., Callea, F., Groote, D. J., Gudat, F., Denk, H., Desmet, V., Korb, G., R.N.M., M., Phillips, M., Portman, B., H., P., Scheuer, P. J., Schmid, M., and Thaler, H. (1995). Histological grading and staging of chronic hepatitis. *Journal of Hepatology*, 22:696–699.

Jähne, B. (1993). *Spatio-Temporal Image Processing: Theory and Scientific Applications*, volume 751 of *Lecture Notes in Computer Science*, chapter Tensor Methods. Springer-Verlag.

Jaster, R., Sparmann, G., and Emmrich, J. (2002). Extracellular signal regulated kinases are key mediators of mitogenic signals in rat pancreatic stellate cells. *Gut*, 51:579–84.

Johne, B. (1993). *Spatio-temporal image processing*. Springer.

Johnson, R. (2005). *Probability and Statistics for Engineers*. Pearson Education International.

Jones, H. A. (2005). Inflammation imaging. *Proc. Am. Thor. Soc.*, 2:545–48.

Joshi, N. (2007). *Non-parametric probability density function estimation for medical images*. PhD thesis, University of Oxford, Department of Engineering Science.

Kadir, T. (2002). *Scale, Saliency and Scene Description*. PhD thesis, University of Oxford.

Kadir, T. and Brady, M. (2001). Scale, saliency and image descriptors. *IJCV*, 45:83–105.

Kalkan, S., Calow, D.and Wörgötter, F., Lappe, M., and Krüger, N. (2005). Local image structures and optic flow estimation. *Network*, pages 341–56.

Kalkan, S., Wörgötter, F., and Krüger, N. (2007). First-order and second-order statistical analysis of 3D and 2D image structure. *Network*, pages 129–60.

KIMIA-database (2011). KIMIA binary image database: Collection I.

Kovesi, P. (1999). Image features from phase congruency. *VIDERE: Journal of Computer Vision Research*, 1(3):1–27.

Krieger, G. and Zetzsche, C. (1996). Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE TIP*, 5:1026–1042.

Krüger, N. and Felsberg, M. (2003). A continuous formulation of intrinsic dimension. In *British Machine Vision Conference*.

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

Lee, W.-L., Chen, Y.-C., and Hsieh, K.-S. (2003). Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform. *IEEE TMI*, 22(3):382–392.

Lefkovicz, L. (2013). Mri galery. `http://www.howstuffworks.com/mri.htm`. [Accessed on 19-09-2013].

Leung, T. and Malik, J. (2001). Representing and recognizing the visual appearance of materials using three-dimensional textons. *Int. J. of Comp. Vis.*, 43:29–44.

Lindeberg, T. (1996). Scale-space: A framework for handling image structures at multiple scales. In *In: Proc. CERN School of Computing, Egmond aan Zee, The Netherlands, 821 September, 1996.*

Macarini, L. and Stoppino, L. (2013). *Radiologic assessment of liver fibrosis – present and future*, chapter Practical management of viral hepatitis. InTech.

Marr, D. (1982). *Vision: A computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: W.H.Freeman.

Marroquin, J., Santana, E., and Botello, S. (2003). Hidden Markov Measure Field models for image segmentation. *IEEE PAMI*, 25:1380–1387.

Mathworld, W. (2014). Spherical coordinates.

Mazoff, C. (2010). HCV diagnostic tools: Grading and staging liver biopsy. *A publication of the Hepatitis C support project*, 2.4:1–2.

MBBS Humanity First, S. M. (2011). Liver ultrasound: Sonographic anatomy and normal findings. `http://medicinembbs.blogspot.co.uk/2011/08/liver-ultrasound.html`. [Accessed on 19-09-2013].

McCarrol, J., Phillips, P., Kumar, R., Park, S., Pirola, R., Wilson, J., and Apte, M. (2004). Pancreatic stellate cell migration: role of phoshpatidylinositol 3-kinease pathway. *Biochem. Pharmacol.*, 67:1215–25.

Mellor, M., Hong, B. W., and Brady, M. (2008). Locally rotation, contrast, and scale invariant descriptors for texture analysis. *IEEE PAMI*, 30(1):52–61.

Morrone, C. and Burr, D. C. (1988). Detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society London, Series B, Biological sciences*, 235(1280):221–245.

Morrone, C. and Owens, R. (1987). Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313.

Morrone, C., Ross, O., Burr, D., and Owens, R. (1986). Mach bands are phase dependent. *Letters to Nature*, 324:250–253.

Mrabet, Y. (2013). For information on multiple sclerosis. `http://www.my-ms.org/mri_plane_math.htm`. [Accessed on 10-09-2013].

Nabighian, M. (1984). Towards a three-dimensional automatic interpretation of potential field data via generalized Hilbert transforms: Fundamental relations. *Geophysics*, 49:780–786.

NessAiver, M. (1997). *All you really need to know about MRI physics*. Harbor Duvall Graphics.

Noterdaeme, O. and Brady, M. (2008). A fast method for computing and correcting intensity inhomogeneities in MRI. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1525–1528.

Oppenheim, A. and Lim, J. (1981). The importance of phase in signals. In *Proceedins of the IEEE*, pages 529–550.

Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of mathematical statistics*, 33(3):1065–1076.

Pennec, X. (2006). Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J Math Imaging Vis*, 25:127–54.

Pennec, X. and Ayache, N. (1998). Uniform distribution, distance and expectation problems for geometric features processing. *J Math Imaging Vis*, 9:49–67.

Pennec, X. and Thirion, J.-P. (1997). A framework for uncertainty and and validation of 3-d registration methods based on points and frames. *IJCV*, 25:203–29.

Perna, A., Tosetti, M., Montanaro, D., and Morrone, M. C. (2008). BOLD response to spatial phase congruency in human brain. *Journal of Vision*, 8:15:1–15.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE TPAMI*, 12(7):629–639.

Perwass, C. (2009). *Geometric Algebra with Applications in Engineering*. Springer.

Pitteway, M. (1967). Algorithm for drawing ellipses or hyperbolae with a digital plotter. *The Computer Journal*, 10(3):282–289.

Randen, T. and Husoy, J. (1999). Filtering for texture classification: a comparative study. *IEEE TPAMI*, 21(4):291–310.

Ratziu, V., Charlotte, F., Heurtier, A., Gombert, S., Giral, P., Bruckert, E., Grimaldi, A., Capron, F., and Poynard, T. (2005). Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *New England Journal of Medicine*, 128:1898–1906.

Reisfeld, D. (1996). The Constrained Phase Congruency feature detector: simultaneous localization, classification, and scale determination. *Pattern Recognition Letter*, 17:1161–69.

Schenk, V. and Sir Brady, M. (2003). Improving phase-congruency based feature detection through automatic scale-selection. *LNCS:CIARP*, 2905:121–8.

Selesnick, I., Baraniuk, R. G., and Kingsbury, N. G. (2005). The Dual-Tree Complex Wavelet Transform. *IEEE SP Magazine*, pages 123–154.

Shao, L. (2005). *Invariant, Salient Regions Based Image Retrieval*. PhD thesis, University of Oxford.

Shi, K. and Zhu, S.-C. (2007). Mapping natural image patches by explicit and implicit manifolds. In *CVPR*.

Signore, A., Mather, S., Piaggio, G., Malviya, G., and Dierckx, R. (2010). Molecular imaging of inflammation/infection: nuclear medicine and optical imaging agents and methods. *Chemical Review*, 110:3112–45.

Signore, A., Procaccini, E., Annovazzi, A., Chianelli, M., van der Laken, C., and Mire-Sluis, A. (2000). The developing role of cytokines for imaging inflammation and infection. *Cytokine*, 12:1445–54.

Smith, S. (1992). *Feature based image sequence understanding*. PhD thesis, University of Oxford, UK.

Smith, S. and Brady, J. (1997). SUSAN - A new approach to low-level image processing. *IJCV*, 23:45–78.

Starkel, P. and Leclercq, I. (2011). Animal models for the study of hepatic fibrosis. *Best Practice and Research Clinical Gastroenterology*, 25:319–33.

Stein, E. (1970). *Singular integrals and differentiability properties of functions*. Princeton University Press.

Styner, M., Brechbuehler, C., Székely, G., and Gerig, G. (2000). Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE TMI*, 19:3:1–14.

Talwalkar, J., Yin, M., Fidler, J., Sanderson, S., and Kamath, P. (2008). Magnetic Resonance Imaging of hepatic fibrosis: emerging clinical applications. *Hepatology*, 47:332–42.

Taouli, B., Tolia, A., Losada, M., Babb, J., Chan, E., Bannan, M., and Tobias, H. (2007). Diffusion-weighted MRI for quantification of liver fibrosis: preliminary experience. *Abdominal Imaging*, 189:799–806.

Tiderius, C., Svensson, J., Leader, P., Ola, T., and Dahlberg, L. (2004). dGEMRIC (delayed Gadolinium-enhanced MRI of cartilage) indicates adaptive capacity of human knee cartilage. *Magnetic Resonance in Medicine*, 51:286–90.

Wang, L. and He, D.-C. (1990). Texture classification using texture spectrum. *Pattern Recognition*, 23:905–910.

Weickert, J. and Scharr, H. (2002). A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *J. Vis. Command. Im. Repr.*, 13:103–18.

Wells, W., Grimson, W., Kikinis, R., and Jolesz, F. (1996). Adaptive segmentation of mri data. *IEEE TMI*, 15:429–442.

Wietzke, L. and Sommer, G. (2008). The 2d analytic signal. Technical report, Christian-Albrechts-Universitat, Zu Kiel.

Wietzke, L., Sommer, G., and Fleischmann, O. (2009). The geometry of 2D image signals. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1690–97.

Wong, A., Clausi, D., and Fieguth, P. (2010). Cpol: Complex phase order likelihood as a similarity measure for mr-ct registration. *Medical Image Analysis*, 14(1):50–7.

Zetzsche, C. and Barth, E. (1990). Fundamental limits of linear filters in the visual processing of two dimensional signals. *Vision Research*, 30:1111–1117.