NTNU
Innovation and Creativity

Edge-Detection in Signals using the Continuous Wavelet-Transform.

Edge-Detection in Medical UltraSound Images.

Preben Gråberg Nes

Master of Science in Mathematics
Submission date: October 2006
Supervisor: Yurii Lyubarskii, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Preface.

This master-thesis was written at the Faculty of Information Technology, Mathematics and Electrical Engineering (IME) at NTNU in 2005/2006. The thesis represents the end of a 6 year study of mathematics, 3 years at Copenhagen University and 3 years at NTNU. The teaching supervisor is Prof. Yurii Lyubarskii.

The purpose of this thesis is to investigate techniques for detecting edges using the continuous wavelet-transform. To fully understand and take control over each step of the algorithms, I implemented the different techniques in MatLab, rather than using already existing wavelet-toolboxes.

I would in particular like to thank Professor Yurii Lyubarskii for help and guidance throughout the work with the thesis. He has been an excellent motivator and has most kindly provided hints and ideas for improvements, both mathematical and for improvement of scientific expression.

A special thank to the researchers at SINTEF Health Research in Trondheim who have provided me with images used in the analysis. During the work they have also helped to define the targets for the analysis, and to see the problem from an applied viewpoint.

I also would like to thank my girlfriend Linn Grepstad. She has helped me proofreading the thesis, and has come with good ideas of improvement and useful criticism. I would also like to thank her for her general support and helping me to relax in hectic periods.

Trondheim, 13. October 2006

_____

Preben Gråberg Nes

Abstract.

Today, UltraSound (US) images are often used in medical examination and surgery. An improvement of the quality of these US-images will lead to many advantages, which is a big motivation for research on this field. One obstacle in improving the quality of the images is the presence of noise and texture. In order to distinguish this unwanted information from the interesting objects, different techniques can be used. Characteristic features, such as the ability to find vague contours, small objects or edges of small strength, decides if the technique is suitable for analysing noisy signals. This thesis presents different techniques for finding objects in US-images by using the continuous wavelet-transform.

One observation from the analysis is that for edge-detectors using the wavelet-transform at a single scale, there is a compromise between accuracy and reliability. One has to choose between detecting small objects or vague contours. At fine scales one is able to detect small objects, but not objects with a vague contour without including redundant information. At coarse scales one is able to detect vague contours without including redundant information, but one will not detect small objects. The Lipschitz-regularity and the length of a maxima-line in the time-scale plane works well to find the points where the signal changes with a long duration, but is less suitable to find small objects and to remove unwanted information. By using the value of the wavelet-transform at several scales, it is possible to find vague contours in images, small objects, and edges of small strength compared to the strength of the noise. Another important observation from the analysis is that the use of the circumference of objects is appropriate in order to find the most important objects in an image. Using this information has been very useful with respect to the analysis of US-images.

Medical ultra-sound images are in general of varying quality. In addition the quality of a US-image will typically change within the signal, and changes

with respect to the quality of the contour of objects and the influence of noise. The technique which in general is most reliable and produces the best representations of the US-images analysed in this thesis, uses information about the amplitude of the wavelet-transform both within and across scales, in addition to information about the circumference of the objects. This combined edge-detector is reliable with respect to represent the important objects in the image, and this representation is often easily obtained by the edge-detector.

Contents

List of Tables

List of Figures

Chapter 1

Introduction.

1.1 Introduction.

This thesis is a study of different techniques which can be used to detect edges in 1-D and 2-D signals. The report will focus on difficulties which occur with respect to noisy signals, in particular to medical US-images. This thesis presents theory of 1-D and 2-D continuous wavelet transform, and which information that can be used for edge-detection purposes. By implementing the different edge-detectors in MatLab and applying them to US-images, strengths and weaknesses of the edge-detectors will be discussed.

The master-thesis is divided into two main-parts:

- Part I: A theoretic study of properties of the 1-D and 2-D continuous wavelet-transform.
- Part II: Study of 1-D and 2-D edge-detectors.

Part I begins with presenting theory regarding the 1-D continuous wavelet-transform. The study begins with defining the wavelet-transform, and illustrating the definition and some useful properties. Properties regarding the decay of the wavelet-transform will be discussed, and illustrated with an example. The second chapter of Part I contains a study of the 2-D continuous wavelet-transform, and some of its properties. Part I ends with a brief discussion regarding noise, and the uncertainty-principle of edge-detection in noisy signals.

Part II begins with a discussing how one may identify and relate the continuous wavelet-transform at different scales. This discussion is one of

the fundamental issues with respect to the analysis of this thesis. The tools which are discussed in this chapter, enables the construction of multi-scale edge-detectors. The discussion and performance of edge-detectors for US-images rely on the ability to combine information at several scales.

The last two chapters of Part II discuss 1-D and 2-D edge-detectors. In these chapters it will be discussed how one can combine properties of the wavelet-transform in order to find the important objects in a signal. The edge-detectors have been implemented in MatLab, and based on the output some characteristic features of each of the edge-detectors will be identified. By comparing with preliminary targets for the US-image, it will be discussed which edge-detector is the most suitable for the type of US-images analysed in this thesis.

1.2 Preliminary Targets and Basic Concepts.

This section serves a two-folded purpose. First, introduce the reader to underlying ideas of some fundamental concepts of this thesis. These concepts are elaborately discussed later in the thesis. The second purpose of this section is to define preliminary targets for the analysis. A survey of what one may hope to achieve at the end of analysis is discussed. These targets are defined from an applied viewpoint, i.e. with respect to what is desirable for use in medical applications.

Sect.(1.2.1) introduces the concepts of *edge* and *edge-detector*. The discussion is not of theoretic disposition, but uses figures and examples to help the reader to understand the underlying ideas of the concepts. With the help of examples it will be explained that there exists no optimal edge-detector for all applications. An edge-detector demand collaboration with an operator. Applications of edge-detection theory will be discussed, in particular applications important in the thesis.

Sect.(1.2.2) discusses preliminary targets for the analysis. The primary target is to analyse one particular ultrasound-image. The key-components of the image are identified, and information of interest for medical personnel is discussed. This identification is used to make a brief account of targets for the analysis. These targets will serve the purpose to compare the performance of different edge-detectors which will be discussed in this thesis.

1.2.1 Edge and Edge-Detector.

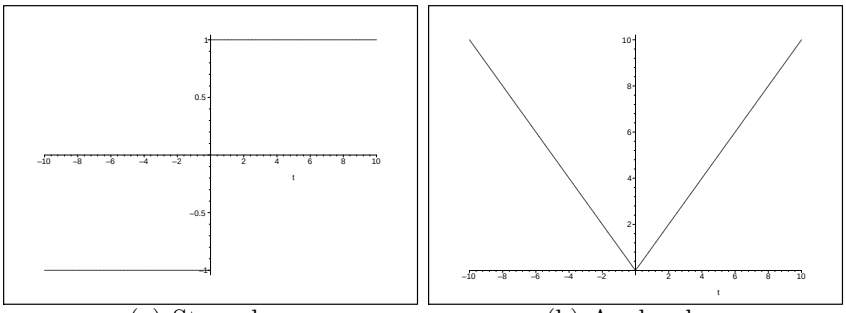
In this section some fundamental concepts of the thesis will be introduced. The concepts *edge* and *edge-detector* will be presented.

The discussion of edge is of intuitive disposition and is included to help the reader to grasp the most important concept of the thesis. Some examples of edges are used to visualize the concept. In Ch.(2) the theoretic definition of edge will be discussed.

Another important concept of the thesis is *edge-detector*. An edge-detector is a machinery constituded by several components. The basic components of an edge-detector will be presented, and some examples of edge-detectors discussed. At the end a short discussion of what one hopes to achieve by using edge-detection theory at medical images.

The principal concept of this thesis is *edge*. When thinking of an edge in

a function one is most likely to think of a point where the value of the function changes rapidly. A point where the amplitude of the function changes



(a) Step-edge. (b) Angle-edge.

Figure 1.1: Illustration of two different kinds of edges.

rapidly is an edge, and denoted *step-edge*. An example of a step-edge is displayed in Fig.(1.1a). Step-edges are typically the most important edges and are the edges one usually seeks in a function or signal. There exists other less obvious kind of edges. If the first derivative of the function has a step-edge at a point, the function has an edge at the point. One denotes the edge by *angle-edge*. An example of an angle-edge is displayed in Fig.(1.1b). In particular step-edges are important in this thesis, but angle-edges will be encountered. Generally, a point where the function or some derivative of the function has a step-edge is called an edge. The theoretic definition of edge is elaborately discussed in Sect.(2.3).

A function typically consists of several edges. An example of a function with several step-edges is presented in Fig.(1.3). When analysing a function all edges are usually not of interest. An edge-detector tries to find the edges in the signal which is of interest for the operator. An edge-detector is not a single component, but rather a machinery composed of several components. A schematic overview of the basic and necessary components of an edge-detector is illustrated in Fig.(1.2). The basis for edge-detectors is the function. The first step of detecting the important edges is to find all the edges in the function. *For the signal in Fig.(1.3) this is all the points marked by a dot.* The component detecting all the edges will in this thesis be refered to as the *trivial edge-detector*. The wavelet based trivial edge-detector will be introduced in Sect.(2.3) and Sect.(3.2). The trivial edge-detector does not take into account the importance of the edges. The set of all edges in the

Figure 1.2: A schematic overview of the basic components of edge-detection. The first step in detecting the important edges in a signal is to find the set of all edges in the signal. An edge-detector uses additional information in order to locate the important edges in the signal.

signal is typically redundant and contains a lot of uninteresting information. The second step of detecting edges in a signal is to use some filter on the set of all edges to separate the important edges from the redundant edges. This filter may be composed of several components. In Ch.(6) and Ch.(7) filters separating unwanted from wanted information will be discussed. The output of the filter should be the information in the signal being of interest for the operator. It is not yet discussed what is meant by *information of interest*. This depends on the actual application. The term *edge-detector* will in this thesis both to refer to the filter removing redundant information and the overall process. The next paragraph discusses examples of edge-detectors.

Consider the signal in Fig.(1.3). The signal have 20 step-edges all marked

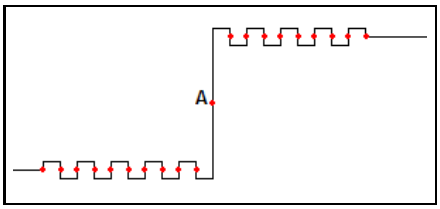


Figure 1.3: A signal with 20 step-edges.

by a dot. The optimal edge-detector detecting all edges in the signal should detect all the dots. The optimal edge-detector detecting the edges where the overall intensity changes should only detect the step-edge marked by **A**. If one is interested in the influence of noise in the signal, the optimal edge-detector should detect all the step-edges except the step-edge marked **A**. All the edge-detectors have their applications. This implies that there is no edge-detector optimal for all purposes, but varies with respect to which information is of interest for the operator. This requires that prerequisites

for the detector are defined prior to the investigation. One has to decide what information one would like to obtain from the signal.

An adage claims that an image tells more than a thousand words. In certain contexts it is useful to explain the image by a few words. This requires that the few words describe the important features of the image. One application of edge-detection theory is for reducing the amount of information in the signal. Said equivalently; emphasize the important information in the image. This thesis will in particular focus on edge-detection in US-images. For medical purposes this may help the surgeon to focus on the vital objects in the image when preparing for surgery, or to visualize the objects during closed surgery.

1.2.2 The Image and Preliminary Targets.

In the previous section the necessity of defining preliminary targets for the edge-detector prior to analysis was discussed. In this section the prerequisites for the image in Fig.(1.4) will be discussed. The targets presented in this section are used to compare edge-detectors, and to see how close the edge-detectors are to being optimal for this particular image.

Prior to discussing the preliminary targets for the edge-detector, the objects in the image are identified. The objects medical personnel would like to get information about are discussed. These objects will be used to define the targets the optimal edge-detector should satisfy. The primary target for the analysis is to find an edge-detector good for the ultrasound-image in Fig.(1.4).

The solid grey-white object in the centre of the image is a brain tumour. The diameter of the brain-tumour is approximately 40 mm. The tumour is the key-component of the image. This object is of particular interest for medical personnel. The information of interest is the size, its form and relation with other objects.

Inside the tumour there are several cysts. From Fig.(1.4) it is difficult to determine the number. Medical personnel are interested in knowing the existence of cysts and their position within the tumour. The actual number of cysts is not of importance.

The horizontal white line visual in the middle-left of the image is the cortex. One can in addition observe a fragment of the cortex to the lower-right of the tumour. This object is not of interest for medical personnel.

Figure 1.4: The US-image is the foundation of the analysis of edge-detectors. The US-image displays a brain-tumour(solid white). Inside the tumour there are small cysts. The horizontal white line is the cortex.

The second topic of this section is to use the identification of the image to define preliminary targets for the edge-detector, i.e. define what information one in advance would like the edge-detector find after analysis. Fig.(1.5) display the US-image with indications of the vital information in the image.

The line marked **A** is the upper half of the contour of the brain-tumour. The transition between the different tissues in this region is "rather" sharp. The line marked **A** should be detected.

The area marked **C** is the lower half of the contour of the brain-tumour. One problem with US-images is that the details decrease away from the US-probe. In addition US-images are vulnerable for shadows. This phenomenon is causing the blurred edge in the region marked by **C**. This cause uncertainty on the actual position of the boarder in this region. One proposed line is marked by the yellow-dotted line. From a medical viewpoint it is sufficient to detect "something" similar to the dotted line. One may observe that the fragment corresponding to the cortex on the right is not marked as interesting.

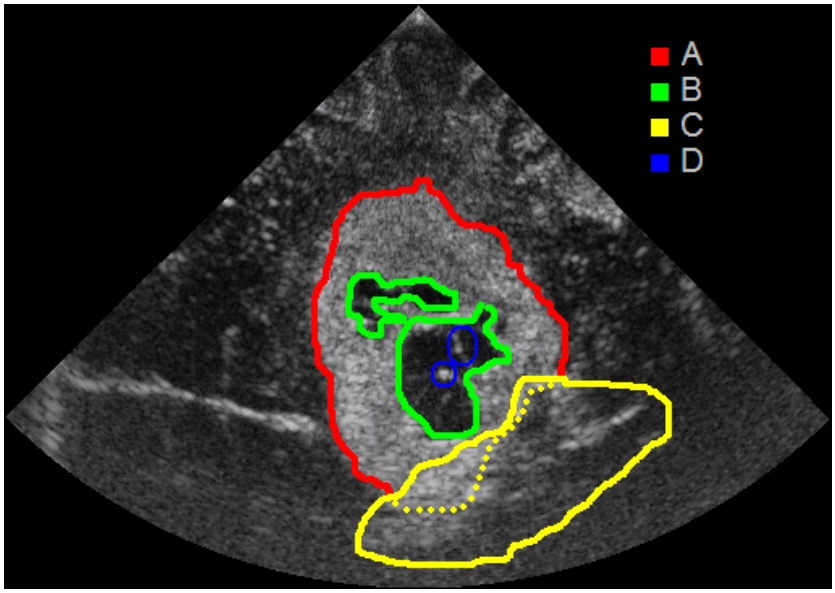


Figure 1.5: The US-image together with a description of the important objects in the US-image. The region marked **A** and **B** should be detected. One might expect more problems in the region marked as **C**. Some small objects are marked by **D**.

The line marked **B** is the contour of the cyst(s). The uncertainty of the number of cysts is discussed. From a medical viewpoint it is not interesting to know the exact number. It is sufficient to consider the collection of cysts as two distinct cysts and find the separation between these. Fig.(1.4) indicate that the contour of the cysts is in some areas is curly. In particular the boarder of the upper cyst is curly. From a medical viewpoint detailed information of the boarder is not of interest. A sketch of the contour is sufficient. A boarder similar to the proposed marked **B** is sufficient for medical purposes. From a theoretic viewpoint it is interesting to see how much information of the contour is possible to achieve.

The regions marked **D** is not of interest from a medical viewpoint. The objects marked will serve to investigate the performance of detecting small fragments in the image.

From a medical viewpoint the following properties summarizes the optimal edge-detector.

- Detect the contour marked as **A**, **B** and the yellow-dotted line.
- Detect as few additional edges as possible.

The first remark is a minimum requirement. One will not be satisfied if any of the edges **A**, **B** or something similar to the yellow dotted-line are undetected. The information obtained from these edges is what being of interest for medical personnel. Detecting these edges is therefore an optimal requirement for the edge-detector. One will in addition most likely detect additional edges. Both the cortex, the outer boarder of the image and "something" in the upper-right will most likely be detected. The intention with the second remark is to detect as few edges not corresponding to an object as possible.

From a theoretic viewpoint the following objects will be of interest.

- Detect the objects marked **D**.
- The contour of the upper cyst(s). It is interesting to see how much detail is possible to achieve.
- The position of the contour corresponding to the yellow-dotted line.

These items are used to test the reliability of edge-detectors at high-details. Several edge-detectors will give sufficient output with respect to information desired from a medical viewpoint.

Part I

Wavelet-Theory.

Chapter 2

1-D Wavelet-Transform.

In this chapter the theory of 1-D continuous wavelet-transform and its properties will be discussed.

The target of Sect.(2.1) is to define wavelets and the 1-D continuous wavelet-transform. In Sect.(2.3) it will be attempted to illustrate why and how edges may be found by the continuous wavelet-transform. This will result in a definition of the wavelet-based trivial edge-detector. As indicated in the previous chapter, the output of the trivial edge-detector typically contains a lot of unwanted information. In Sect.(2.5) one of the important characteristics which may be used to find the important edges in a function will be introduced, the *maxima-tree*.

One remarkable property of the wavelet-transform is that it is possible to extract information about which kind of edges one is dealing with. One may determine if an edge is e.g. a step-edge or an angle-edge. The concept which makes this possible is the Lipschitz-regularity, and will be defined in Sect.(2.2). The relation between Lipschitz-regularity and different kind of edges is discussed in Sect.(2.3), and how one may extract this information is discussed in Sect.(2.4).

In Sect.(2.6) some obstacles in estimating the Lipschitz-regularity of an edge in numerical computations will be illustrated. This discussion will result in various estimates of how one may improve the estimate of the Lipschitz-regularity. In Sect.(2.7) it will be attempted to estimate the Lipschitz-regularity of some edges in a signal.

2.1 1-D Wavelet-Transform.

In this section the topics of *wavelet* and *continuous wavelet-transform* will be presented. These topics will be presented with the help of underlying ideas and some examples to help the reader grasp some important properties.

In Sect.(1.2.1) the concept *trivial edge-detector* was introduced as the first step for identifying the edges of a signal. In this thesis the continuous wavelet-transform will be used as the basis of the trivial edge-detector. The *wavelet-based trivial edge-detector* for 1-D signals will be defined in Sect.(2.3). The CWT contains additional information which may be used to detect the important edges in a signal. How this information can be obtained and how it can be used will be discussed in Sect.(2.4) and Ch.(6).

Gaussian derived wavelets and the wavelet-transform with respect to these wavelets will be important in the thesis. Examples of these and their wavelet-transform are calculated for some signals. It will be proved that the *amplitude* of step-edges are in some sense reflected by the wavelet-transform.

The section begins with definition and examples of wavelets followed by definition and examples of the continuous wavelet-transform. Some important properties of the concepts will be discussed.

First the definition of wavelet.

Definition 2.1.1 (1-D Wavelet) *A 1-D wavelet is a function $\psi \in L^2(\mathbb{R})$ such that*

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega}\, d\omega < \infty. \tag{2.1.1}$$

Eq.(2.1.1) is known as the "admissibility condition". $\hat{\psi}$ is the Fourier-transform of ψ. Let

$$\psi_{u,s} = \frac{1}{\sqrt{s}}\psi\Big(\frac{t-u}{s}\Big), \tag{2.1.2}$$

denote the scaled and translated wavelet. Denote $\psi_{0,s} = \psi_s$.

□

By a change of variable one can prove that $\psi_{u,s}$ satisfies the admissibility-condition if and only if ψ satisfy the admissibility-condition. One requirement for the integral in Eq.(2.1.1) to be finite is that $\hat{\psi}(0) = 0$. Otherwise the

integral is divergent. By definition of Fourier-transform (see Def.(B.0.1)), this implies that;

$$\hat{\psi}(0) = 0 \Leftrightarrow \int_{-\infty}^{\infty} \psi(t)\, dt = 0, \tag{2.1.3}$$

i.e. wavelets have zero-mean. It is not sufficient that $\hat{\psi}(0) = 0$, but the zero-mean property is important for detecting edges in a signal. This characteristic will help to ensure that points not being an edge are not falsely detected. Detecting edges in a signal is often done in correlation with removing noise in a signal, i.e. to find a "noise-free" version of the function or signal. This requires the existence of an inverse wavelet-transform. The inverse is only defined if the admissibility condition is satisfied, justifying the definition of wavelets by the admissibility-condition[1]. A proof of the inverse CWT can be found in [1].

Example 2.1.1 *The "original" wavelet is the Haar-wavelet defined as;*

$$\psi(t) = \begin{cases} 0 & \text{for } t \notin]-1, 1[\\ -1 & \text{for } t \in [-1, 0] \\ 1 & \text{for } t \in]0, 1] \end{cases}$$

The Haar-wavelet trivially fulfil the latter requirement for a wavelet.
An important class of wavelets is the wavelets corresponding to (the normalized) Gaussian. They are given by;

$$\psi(t) = (-1)^n \frac{d^n}{dt^n}\theta(t),$$

for $n \in \mathbb{N}_+$, *where*

$$\theta(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \in C^\infty(\mathbb{R}).$$

By rules of Fourier-transform of differentiation and Fourier-transform of the Gaussian (see Prop.(B.0.1)) it follows;

$$\hat{\theta}^{(n)}(\omega) = (i\omega)^n \hat{\theta}(\omega) = (i\omega)^n e^{-\frac{1}{2}\omega^2}.$$

Since $\omega^n e^{-\frac{1}{2}\omega^2} \in L^2(\mathbb{R})$ *for all* $n \in \mathbb{N}$ *the admissibility-condition is fulfilled. Two important examples of Gaussian derived wavelets are the 1.st and 2.nd*

[1] In a lot of edge-detection literature the admissibility-condition is not used to define wavelets, only that $\hat{\psi}(0) = 0$.

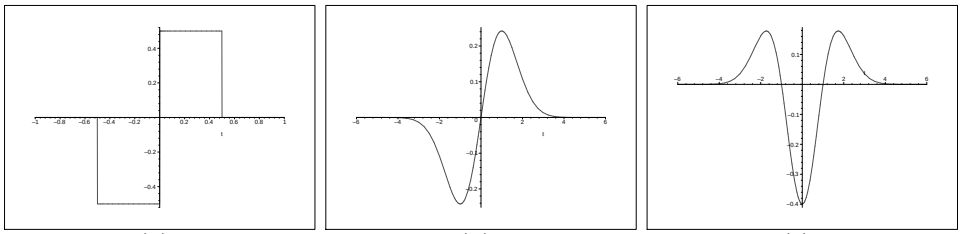
derivative, given by;

$$\psi^1(t) = \frac{1}{\sqrt{2\pi}} t e^{-\frac{t^2}{2}} \tag{2.1.4}$$

$$\psi^2(t) = \frac{1}{\sqrt{2\pi}} (t^2 - 1) e^{-\frac{t^2}{2}} \tag{2.1.5}$$

The wavelet in Eq.(2.1.5) is known as The Mexican Hat. Fig.(2.1) display the wavelets. ψ^1 will be used to denote the 1-D wavelet corresponding to the 1.st derivative of the Gaussian, and ψ^2 will be used to denote the 1-D wavelet corresponding to the 2.nd derivative of the Gaussian.

□



(a) (b) (c)

Figure 2.1: (a) The Haar-wavelet. (b) The wavelet corresponding to the 1st derivative of the Gaussian-function. (c) The wavelet corresponding to the 2nd derivative of the Gaussian-function, ofted denoted *The Mexican Hat*.

By Ex.(2.1.1) it follows that there exists more than one wavelet. In Sect.(2.3) it will be explained which wavelets are suitable for detecting different kind of edges in a signal.

The next step is to define the *continuous wavelet-transform* (CWT). Several definitions may be found in literature. The definition used in this thesis is by Mallat in [1][2].

Definition 2.1.2 (Continuous Wavelet-Transform.) *Let $f \in L^2(\mathbb{R})$. The continuous wavelet-transform is defined as;*

$$Wf(u,s) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \Big(\frac{t-u}{s} \Big) \, dt, \tag{2.1.6}$$

[2]Mallat uses several definitions of the wavelet-transform. Be sure to use [1] as a reference.

*where * represents the complex conjugate.*

□

In some applications it can be useful to represent the wavelet-transform as a convolution-product. Denote $\tilde{f}(t) = f(-t)$. Then;

$$Wf(u,s) = f * \tilde{\psi}_s(u). \tag{2.1.7}$$

Eq.(2.1.7) is particular useful for the Gaussian derived wavelets. Assume $\psi(t) = (-1)^n \theta^{(n)}(t)$. Pr.(B.0.2) relates the convolution and differentiation. The wavelet-transform with respect to this wavelet can be expressed as

$$Wf(u,s) = f * s^n \frac{d^n}{du^n}\theta_s(u) = s^n \frac{d^n}{du^n}(f * \theta_s)(u). \tag{2.1.8}$$

The expression in Eq.(2.1.8) is fundamental of this thesis, and is used throughout the thesis.
Note: The CWT can be extended to tempered distributions such as the Dirac $\delta(t)$ and the Heaviside-function $u(t)$.

The next step is to compute the wavelet-transform for some examples.

Example 2.1.2 *Let $f(t) = Au(t)$ where $u(t)$ is the Heaviside-function, and $A \in \mathbb{R}$. Assume $\psi^1(t) = -\frac{d}{dt}\theta(t)$. Then*

$$\begin{aligned} Wf(u,s) &= \int_0^\infty \frac{A}{\sqrt{s}}\psi\big(\frac{t-u}{s}\big)\, dt = \frac{As}{\sqrt{s}}\int_{-u/s}^\infty \psi(y)\, dy = \frac{As}{\sqrt{s}}\int_{-u/s}^\infty \frac{-d}{dy}\theta(y)\, dy \\ &= As\tilde{\theta}_s(u) = As\theta_s(u). \end{aligned} \tag{2.1.9}$$

The last equality follows by symmetry of the Gaussian and the second equality follows by a change of variable $sy = t - u$.
Let $f(t) = \delta(t)$ be the Dirac. The wavelet-transform of the Dirac is given by;

$$Wf(u,s) = \delta * \tilde{\psi}^1_s(u) = \tilde{\psi}^1_s(u) = -\psi^1_s(u). \tag{2.1.10}$$

□

From Ex.(2.1.2) the following observation;

Statement 2.1.1 *Assume $|A| > |B|$ for $A, B \in \mathbb{R}$, and $f(t) = Au(t)$, $g(t) = Bu(t)$. Then for all $s > 0$,*

$$\sup_{u\in\mathbb{R}} |Wf(u,s)| = |Wf(0,s)| \geq \sup_{u\in\mathbb{R}} |Wg(u,s)| = |Wg(0,s)|. \tag{2.1.11}$$

□

This implies that the CWT has larger value for large step-edges than small step-edges. In Ch.(6) and Ch.(7) edge-detectors using this fact will be discussed.

2.2 Lipschitz-Regularity.

In this section the *Lipschitz-regularity* of a function will be defined. The concept will be illustrated by the help of some properties and examples.

As discussed in Sect.(1.2.1), the notion of *edge* is closely related to continuity of the function. Lipschitz-regularity is a *quantification* of the continuity of a function. In Sect.(2.3) the relation between Lipschitz-regularity and edge will be established. One reason for using Lipschitz-regularity is that the wavelet-transform preserves information about the regularity of the function. This remarkable property and its consequences will be discussed in Sect.(2.4).

As indicated in Sect.(1.2.1) step-edges and angle-edges are important in this thesis. The Lipschitz-regularity of these edges will be calculated. The relation between the continuity of a function and the Lipschitz-regularity will be discussed.

The discussion of Lipschitz-regularity begins with definition and properties. The definition of Lipschitz-regularity will be extended to distributions, before an example is calculated demonstrating the concept.

Definition 2.2.1 (Lipschitz Conditions, Regularity and Singularity [1])

Local Lipschitz *Let $n \leq \alpha < n+1$ for $n \in \mathbb{N}$. A function $f(t)$ is said to be pointwise Lipschitz-α at t_0 if and only if there exists an $A \in \mathbb{R}_+$, an $\epsilon > 0$ and a polynomial $P_n(t)$ of degree n such that for all $|h| < \epsilon$*

$$|f(t_0+h) - P_n(h)| \leq A|h|^\alpha$$

Global Lipschitz *Let $n \leq \alpha < n+1$ for $n \in \mathbb{N}$, and $]a,b[$ be an open interval. Then $f(t)$ is said to be globally Lipschitz-α over $]a,b[$ if and only if there exists an $A \in \mathbb{R}$ and $\epsilon > 0$ such that for any $t_0 \in]a,b[$ there exists a polynomial P_n of degree n such that if $t_0 + h \in]a,b[$*

$$|f(t_0+h) - P_n(h)| \leq A|h|^\alpha$$

Lipschitz Regularity *The Lipschitz-regularity of $f(t)$ locally at t_0 or globally over $]a,b[$ is the superior bound of all α-values such that $f(t)$ is Lipschitz-α at t_0 or over $]a,b[$.*

Singular *$f(t)$ is said to be singular at t_0, if it is not Lipschitz 1 at t_0.*

□

Note that if a function f is globally Lipschitz-α over $]a,b[$, then f is at least locally Lipschitz-α at any $t_0 \in]a,b[$. Calculating the Lipschitz-exponent of a function may in practice be a troublesome task. The following properties may simplify the task of determining the Lipschitz-regularity of a function.

Property 2.2.1 (Continuity) *Assume $f : \mathbb{R} \to \mathbb{R}$ is n-times continuously differentiable in a neighbourhood of t_0. Then $f(t)$ is Lipschitz-n at t_0.*

Proof: Assume $f : \mathbb{R} \to \mathbb{R}$ is n-times continuously differentiable at t_0. The Taylor-expansion formula gives that in a small neighbourhood of t_0, $f(t)$ can be expressed as

$$f(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(t_0)}{k!}(t-t_0)^k + E_n(t_0) = P_{n-1}(t) + E_n(t_0),$$

with

$$E_n(t_0) \leq \frac{|t-t_0|^n}{n!} \sup_{\nu \in]t_0-\epsilon, t_0+\epsilon[} |f^n(\nu)|.$$

From this it follows that

$$|f(t) - P_{n-1}(t)| \leq |t-t_0|^n K.$$

Def.(2.2.1) give that the Lipschitz-regularity at t_0 is n.

□

Mallat states in [1], [4], [5] and [6] the following three properties of the Lipschitz-regularity.

Property 2.2.2 (Primitive) *If $f(t)$ is Lipschitz-α then its primitive $F(t)$ is Lipschitz-$(\alpha+1)$.*

□

Property 2.2.3 (Derivative) *$f(t)$ is globally Lipschitz-α over $]a,b[$, $\alpha > 1$ if and only if $f'(t)$ is globally Lipschitz-$(\alpha-1)$ over $]a,b[$.*

□

Property 2.2.4 (Continuity) *If $f(t)$ is Lipschitz-α in a neighbourhood of t_0 for $n < \alpha < n+1$ then f is n-times continuously differentiable in a neighbourhood of t_0.*

□

The following property is a warning for relation derivation with Lipschitz.

Property 2.2.5 (Derivation [4]) *If $f(t)$ is Lipschitz-α then $f'(t)$ is not necessarily Lipschitz-$(\alpha-1)$.*

□

A commonly used counter-example proving Prop.(2.2.5) is the function $\sin\frac{1}{t}$ in the neighbourhood of 0.

The Lipschitz-exponent is by definition only defined for positive α. The next definition will extend Lipschitz-exponent to negative values for distributions.

Definition 2.2.2 ([4]) *Let $f \in \mathcal{S}'$, let $\alpha \in \mathbb{R}/\mathbb{Z}$ and $[a,b]$ be an interval of $\mathbb{R}$. Then $f(t)$ is said to be uniformly Lipschitz-α over $]a,b[$ if and only if its primitive $F(t)$ is uniformly Lipschitz-$(\alpha+1)$ over $]a,b[$.*

□

Example 2.2.1 *The Heaviside. For $t \neq 0$ the Heaviside is constant and therefore $C^\infty(]a,b[)$ if $0 \notin]a,b[$. By Prop.(2.2.1), $u(t)$ is Lipschitz-n for all $n \in \mathbb{N}$ for $t \neq 0$. For $t = 0$, $P_0(h) = 0$ and $A \geq 1$ it follows*

$$|u(h)| = u(t) \leq Ah^0 = A. \tag{2.2.1}$$

It follows that the Heaviside is Lipschitz-0 at 0. Assume that $u(t)$ is Lipschitz-α at $t = 0$ for $\alpha > 0$. Let $A \in \mathbb{R}$ be arbitrary and $P_0 = 0$. By choosing $\epsilon \leq \frac{1}{A^{-\alpha}}$, it follows for $0 < h < \epsilon$

$$|u(h)| = 1 < Ah^\alpha < A\epsilon^\alpha < A(\frac{1}{A^{-\alpha}})^\alpha = 1.$$

This contradicts that the Lipschitz-regularity is larger than 0 at $t = 0$. By Def.(2.2.1) the Heaviside has local Lipschitz-regularity 0 at 0, implying that the global Lipschitz-regularity over $\mathbb{R}$ and any interval which contains 0 is 0.

For an angle-edge, for instance $f(t) = |t|$ it follows that for $t \neq 0$ $f(t)$ is Lipschitz-n for all $n \in \mathbb{N}$. At $t = 0$ one see by choosing $A > 2$, $P_1(h) = h$ and $\alpha = 1$

$$|f(h) - P_1(h)| = ||h| - h| = \begin{cases} 0 & \text{for } h \geq 0 \\ 2|h| & \text{for } h < 0 \end{cases} < A|h|^1, \qquad (2.2.2)$$

that $f(t)$ is Lipschitz-1 at 0. Assume that $1 < 1 + \alpha < 2$. If $f(t)$ is locally Lipschitz-$(1+\alpha)$ at 0, it follows that the global Lipschitz-regularity over $\mathbb{R}$ is $1 + \alpha$. Prop.(2.2.3) give that $f'(t) = u(t) - u(-t)$ is globally Lipschitz-$\alpha > 0$ contradicting the fact that $u(t)$ is globally Lipschitz-0 over $\mathbb{R}$.

From the last example it follows that $f(t)$ is globally Lipschitz-α for all $0 < \alpha < 1$. An angle-edge is the 2nd primitive of the Dirac, and by Def.(2.2.2) it follows that a Dirac is Lipschitz-α for $\alpha < -1$.

□

2.3 The Wavelet-Transform and Edges.

In this section an definition of *edge* will be discussed, and how edges can be detected by the wavelet-transform. The relation between edges and the wavelet-transform and the *zooming-property* of wavelets is discussed.

In Sect.(1.2.1) a large class of edges were identified. If a function, in which the function itself or any derivative has a step-edge at a point, it was said that this function has an edge at the point. It was not discussed if these were the only edges. In this section all edges will be identified. This identification is done by using the Lipschitz-regularity defined in the previous section. It will be discussed why some wavelets detects edges in a signal. Wavelets allow one to zoom in or out of the signal. Why one may do this and consequences will be discussed with the help of an example.

We will see that there exists edges of a more peculiar kind than the ones introduced in Sect.(1.2.1), and that wavelet corresponding to the 1.st derivative of the Gaussian is suitable for detecting step-edges.

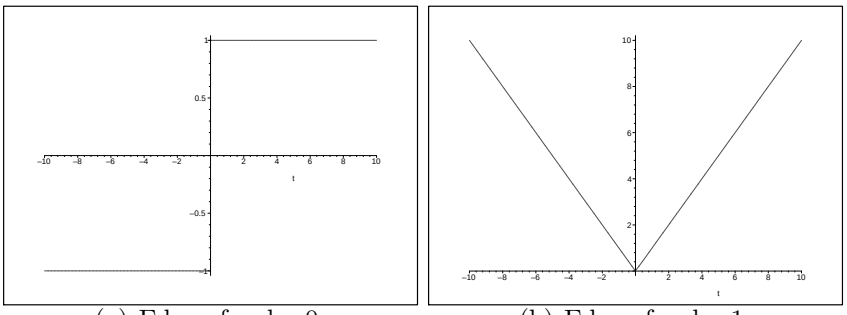
The section begins with a definition of edges. It is discussed that step-edges can be detected by certain wavelets, before introducing the zooming-property of the CWT.

Definition 2.3.1 (Edge) *Assume the function $f(t)$ is Lipschitz-α at t_0 $n \leq \alpha < n+1$, for some $n \in \mathbb{N}$. Then $f(t)$ has an edge of order n at t_0. $(\alpha - n)$ characterises the singularity type of the edge.*

□

The edges introduced in Sect.(1.2.1) all has singularity-type 0.

Example 2.3.1 *The function $f(t) = u(t) - u(-t)$ has an edge of order 0 at 0 see Fig.(2.2). This follows by Ex.(2.2.1) which verified that the Lipschitz-regularity at 0 is 0. The function $g(t) = |t|$ has an edge of order 1 at 0 see Fig(2.2).*



(a) Edge of order 0. (b) Edge of order 1.

Figure 2.2: Edges of order 0 and 1. Edges of order 0 will be called step-edges. Edges of order 1 are called angel-edges.

The steps in the Devils-staircase are edges of order 0. The singularity type of the steps are bounded by $0.44 < \alpha < 0.64$. This is verified in [4].

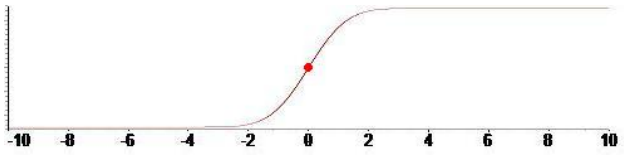
The Dirac has an edge of order -1 at $t_0 = 0$.

□

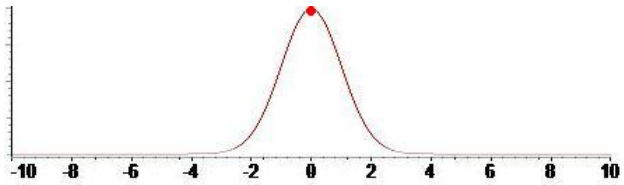
It is clear that there exists other kind of edges that the one discussed in Sect.(1.2.1). These edges are however of little interest for the analysis in this thesis. Three kind of edges are of interest for the purpose of this thesis; Dirac-edges, step-edges and angle-edges. The Dirac-edge which will be used to represent noise, and changes of "short" duration caused by "thin" objects

in the signal. The step-edge will represent sudden changes of "long" duration in the signal.

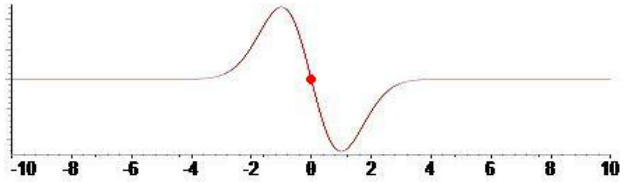
The next question is why and how the wavelet-transform detects edges in a function. To answer these questions, one has to decide in which way one would like the edges to be detected. This is only discussed for the Heaviside function $u(t)$ (step-edges). It will be explained how this can be extended to arbitrary edges of positive order (Primitives of the Dirac) with singularity-type 0. In college-math one learns that the point at which a



(a) $f(t)$.



(b) $f'(t)$.



(c) $f''(t)$.

Figure 2.3: A smooth function together with its 1.st and 2.nd derivative.

smooth function has maximal growth is the point where the derivative is maximal. Equivalently this is the point where the 2.nd derivative has a zero-crossing. Consider Fig.(2.3). Although finding maxima of 1.st derivative is equivalent of finding zero-crossing of 2.nd derivative, the maxima approach will be used in this thesis. One reason is that it is possible to extract useful information of the maxima which is more difficult to extract from the zero-crossings. How this can be done will be discussed in Sect.(2.5). The idea of

finding edges in a signal is to imitate this approach. The problem is that a function with a step-edge is not differentiable at the edge.

Consider the expression of the wavelet-transform for the wavelets corresponding to some derivative of the Gaussian in Eq.(2.1.8);

$$Wf(u,s) = s\frac{d}{du}(f * \tilde{\theta}_s)(u).$$

Assume $f(t)$ has a step-edge at 0. Pr.(B.0.2) verifies that the convolution product $f * \theta_s(t)$ is $C^\infty(\mathbb{R})$, i.e. one can use the above approach to find the points where the derivative of $f * \theta_s(t)$ is maximal, equivalently the points where $Wf(u,s)$ is maximal. One is not guaranteed that the maxima are located at the edge for a general signal. Canny verifies in [3] that the wavelet corresponding to by the 1.st derivative of the Gaussian is close to optimal to ensure good localization for step-edges in a signal. Assume $g(t)$ has an angle-edge at 0. The wavelet-transform with respect to the Mexican-Hat is;

$$Wg(u,s) = s^2\frac{d^2}{du^2}(g * \theta_s)(u) = s^2\frac{d}{dt}(f * \tilde{\theta}_s)(u),$$

by the relation between convolution and derivation, and the since $g'(t) = f(t)$ in sense of distributions. A similar relation holds for the n.th primitive of the Dirac and the $(n+1)$.st derivative of the Gaussian for any $n \in \mathbb{N}$.

There is a second approach to visualise why some wavelets detects step-edges. Fig.(2.4) show a step-edge at 0 and a wavelet at two locations centred at 0 and 10. The wavelet is compactly supported and *odd*[3]. The wavelet-transform at a point u can be considered as the difference between the value of the product $f(t)\psi_s(t-u)$ of the left versus the right side of u. At a point where the function changes little, for instance at $u = 10$ the difference is little giving a small value of the wavelet-transform is small. This is one reason why $\hat{\psi}(0) = 0$ is an important property of wavelets. If however the signal changes value within the support, the difference between the left and right will be larger. Since ψ is odd, this difference will be maximal if the function changes value at the point where the wavelet is centred. The maxima can be located by finding the maxima of the wavelet-transform. Not all wavelets are compactly supported. For instance the wavelets corresponding to the Gaussian are not compactly supported. The energy of these wavelets is concentrated, and a similar argument as above can be used.

[3]The "oddness" of the wavelet is one typical characteristic of wavelets used to detect step-edges. See e.g. the Haar-wavelet, or the spline-wavelet in Ex.(2.5.1).

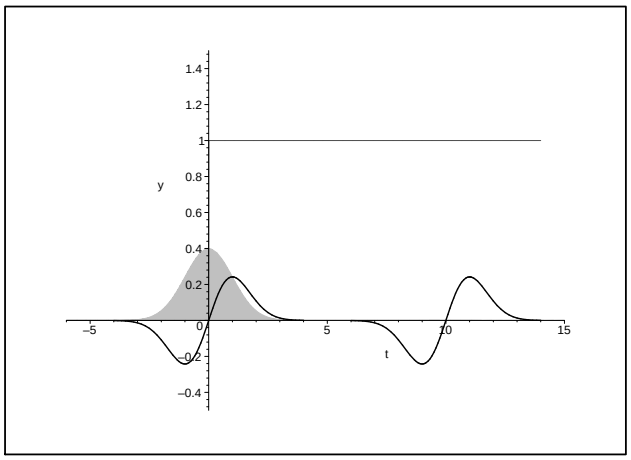
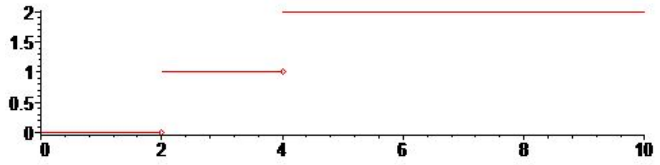
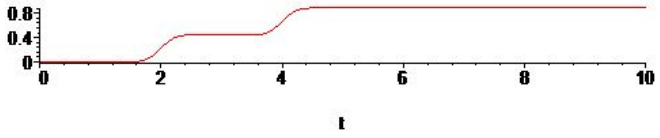


Figure 2.4: Demonstration of why the continuous wavelet-transform detects step-edges.



(a) Two step-edges.



(b) The two step-edges smoothed with the Gaussian-function.

Figure 2.5: Demonstration of why the wavelet-transform detects edges by using convolution.

At this point one may have the impression that a wavelet can only detect edges of one particular kind, e.g. that the wavelet corresponding to the

first derivative of the Gaussian can only detect step-edges, and that the Mexican-Hat only detects angle-edges. If this was the case, one would face an unwanted situation. With respect to the wavelet corresponding to by the first derivative of the Gaussian, Ex.(2.1.2) verify that the response of the wavelet-transform to a Dirac-edge has two maxima. On the other hand one may verify that the response to an angle-edge has no local maxima. Similarly one may verify that with respect to the Mexican-Hat, the response of the wavelet-transform to a step-edge has two maxima, and with respect to a Dirac-edge there are three maxima. It is clear that by using the maxima-approach for detecting edges, one has to chose wavelets with care. One has to chose a wavelet which detects the kind of edges one would like to find, without including too many maximas. In Sect.(2.5) this topic will be further discussed.

To sum up; in this thesis edges are detected by local maxima of the wavelet-transform, denoted *Modulus-Maximum*.

Definition 2.3.2 (Modulus-Maximum) *A modulus-maximum of* $Wf(u,s)$ *is point* u_0 *where* $Wf(u,s)$ *is locally maximum. The maximum should be strict from either left or right.*

□

The last topic which is presented in this section is the *zooming-property* of the continuous wavelet-transform. This is one of the properties which make the wavelet-transform useful for detecting edges. Once again the discussion is restricted to step-edges, and explained by the help of a figure. One may observe that the wavelet-transform uses a variable denoted s. The variable s represents the *scale* of the wavelet-transform, or the "zoom". Signals are usually not as nice as the ones discussed up to this point, but usually contain several edges. Consider Fig.(2.6c) and Fig.(2.6d). Both are the wavelet-transform of the signal in Fig.(2.6a) but with two different scales; a fine/small scale in Fig.(2.6c) and a coarse/large scale in Fig.(2.6d). The zooming property allow to focus in or out on structures in the signal; if one focuses in one the signal one may see that there are two jumps as in Fig.(2.6a), or it one focus out and see the signal from far away only one step is visible as in Fig.(2.6b). At coarse scales one will detect overall changes in the function, while at fine scales one will in addition detect the fine structures. This property is useful for detecting important edges in noisy signals. At coarse scales one will discriminate changes caused by noise and changes of short duration.

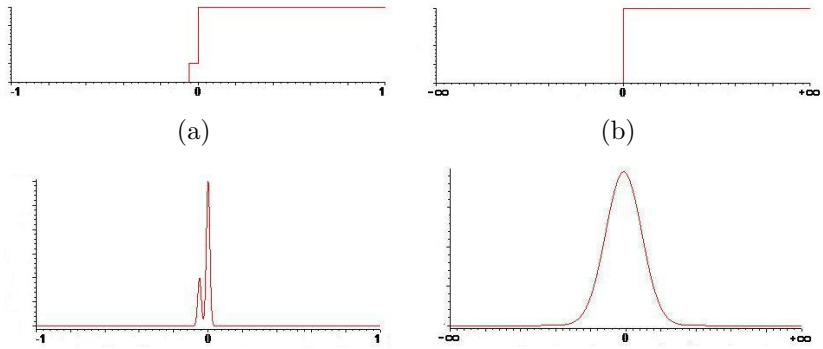
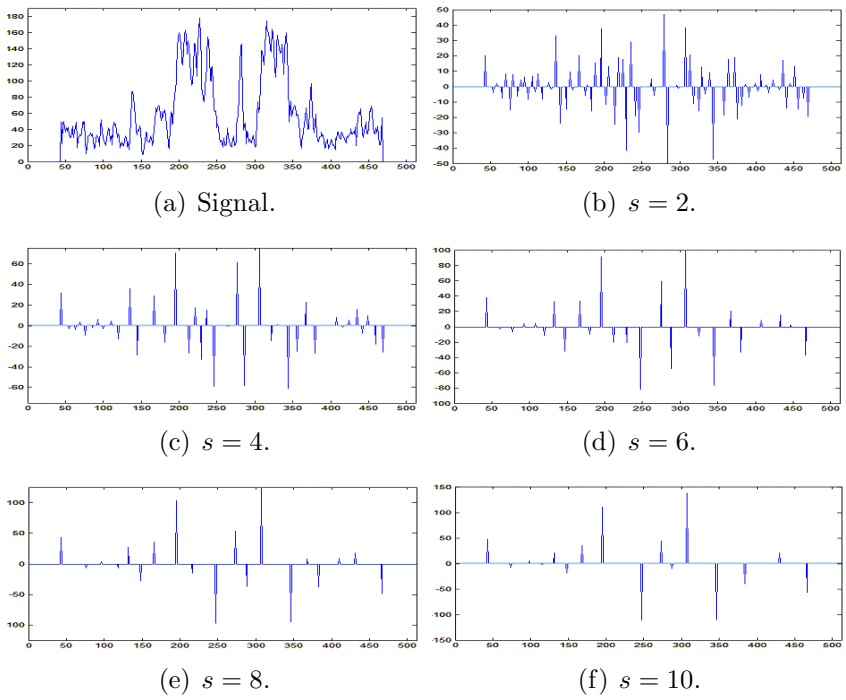


Figure 2.6: Demonstration of the zooming-property of the wavelet-transform.

It will be discussed that edge-detection in noisy signals, such as US-images, would be difficult without the zooming-property of the wavelet-transform.

In Sect.(1.2.1) the first step of detecting the important edges in a signal was to use a trivial edge-detector to detect all the edges. *The 1-D wavelet-based trivial edge-detector will be defined as the set of all modulus-maximum at all the scales the wavelet-transform is computed.*

Fig.(2.7) illustrate the output of the wavelet-based trivial edge-detector at some scales. First one may observe the effect of the zooming-property. At coarse scales there are only a few modulus-maximum, and compared with the signal, these modulus-maximum correspond to the points where the signal has relatively large overall changes. At fine scale there are a large number of modulus-maximum, corresponding to major changes and minor changes in the signal. Another interesting observation is how the value of the modulus-maximum differs across scales. For the "thin" edge or the Dirac-edge in the centre of the image, the values of the modulus-maximum increases as the scale decreases. On the other hand, for the step-edge at 200 the value decreases from approx. 100 at coarse scales to approx. 40 at fine scale. This coincides with the calculations in Ex.(2.1.2). This characteristic will be used in the succeeding chapters to distinguish modulus-maximum corresponding to different kind of edges. The final observation the writer would like to emphasize is the propagation of modulus-maximum across scales. One may observe that modulus-maximum at two different scales apparently reflect the same edge in the signal. This will be discussed in Sect.(2.5) and Sect.(5.1).



(a) Signal. (b) $s = 2$.

(c) $s = 4$. (d) $s = 6$.

(e) $s = 8$. (f) $s = 10$.

Figure 2.7: The output of the wavelet-based trivial edge-detector applied to the signal in Fig.(2.7a).

2.4 The Wavelet-Transform and Lipschitz-Regularity.

One remarkable property of the wavelet-transform is that it contains information about the Lipschitz-regularity of the function. In this section it will be discussed how this information can be obtained from the wavelet-transform.

This property allow one to distinguish different kind of edges in a function, such as Dirac-edges, step-edges and angle-edges with Lipschitz-regularity resp. -1, 0 and 1. This will allow to separate different structures in a function by estimating the Lipschitz-regularity.

Local and global relations between the Lipschitz-regularity and the CWT will be discussed, and that the wavelet corresponding to the n'th derivative of the Gaussian can estimate Lipschitz-regularity for $\alpha < n$.

The section begins with a presentation of some preliminary theory needed to obtain a relation between Lipschitz and CWT, before proving the relations and illustrating the results with some examples.

Before investigating the relation between the CWT and the Lipschitz-regularity, the idea behind the first step of the relation will be presented. A consequence of the definition of Lipschitz-regularity is that there exist an *error-function* denoted $e_\alpha(t)$ and a polynomial $p_m(t)$ of degree m such that;

$$f(t) = p_m(t) + e_\alpha(t), \tag{2.4.1}$$

with

$$e_\alpha(t) \leq K|t - t_0|^\alpha.$$

The underlying idea to estimate the Lipschitz-regularity is to find a wavelet which annihilates the polynomial-term of Eq.(2.4.1), and find some way to measure α from the error term. Before discussing how α can be measured, one need to find wavelets which annihilate polynomial. A wavelet which annihilate a polynomial of degree of order $N-1$, is said to have *N-vanishing moments*.

Definition 2.4.1 (Vanishing moments) *A wavelet $\psi_s(t)$ is said to have N vanishing moments if for all $0 \leq k < N$*

$$\int_{-\infty}^{\infty} t^k \psi_s(t)\, dt = 0$$

□

All wavelets have one vanishing moment since $\hat{\psi}(0) = 0$. By the linearity of the wavelet-transform, wavelets with N vanishing moments is orthogonal to all polynomials $p(t)$ of $deg(p) \leq N-1$. Assume $\psi(t)$ has N-vanishing moments, $N > \alpha$ and $m = \lfloor \alpha \rfloor$. Then;

$$\begin{aligned} Wf(u,s) &= \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi\big(\frac{t-u}{s}\big) dt \\ &= \int_{-\infty}^{\infty} \big(p_m(T) + e_\alpha(t)\big) \frac{1}{\sqrt{s}} \psi\big(\frac{t-u}{s}\big) dt \\ &= \int_{-\infty}^{\infty} e_\alpha(t) \frac{1}{\sqrt{s}} \psi\big(\frac{t-u}{s}\big) dt, \end{aligned}$$

i.e. the wavelet annihilates the polynomial in the representation of $f(t)$.

Definition 2.4.2 (Fast decay) *A function f has fast decay if there for all $m \in \mathbb{N}$ exists a constant $C_m \in \mathbb{R}$ such that*

$$\forall\, t \in \mathbb{R},\ |f(t)| \leq \frac{C_m}{1+|t|^m}$$

□

The question is how one can construct a wavelet with N-vanishing moments. The next theorem will guarantee the existence of wavelets with arbitrary number of vanishing moments, and how these can be constructed.

Theorem 2.4.1 ([1]) *A wavelet $\psi(t)$ with fast decay has n vanishing moments if and only if there exists a $\theta(t)$ with fast decay such that*

$$\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n}$$

Moreover, the wavelet has only n vanishing moments if and only if $\int_{-\infty}^{\infty} \theta_s(t)dt \neq 0$.

□

One example of a function θ satisfying Th.(2.4.1) is the Gaussian, i.e. the wavelet corresponding to the N'th derivative of the Gaussian have N vanishing moments.

The next step is to find the relation between the wavelet-transform and the Lipschitz-exponent. Assume $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz-α at ν. Assume $\psi(t)$ has N vanishing moments with $N > \alpha$. Then;

$$\begin{aligned}
|Wf(u,s)| = |We_\alpha(u,s)| &\leq \int_{-\infty}^{\infty} K|t-\nu|^\alpha \Big|\frac{1}{\sqrt{s}}\psi\big(\frac{t-u}{s}\big)\Big| dt \\
&= K\sqrt{s}\int_{-\infty}^{\infty} |xs+u-\nu|^\alpha |\psi(x)| dx \\
&\leq K\sqrt{s}2^\alpha \Big(\int_{-\infty}^{\infty} |xs|^\alpha |\psi(x)| dx + |u-v|^\alpha \int_{-\infty}^{\infty} |\psi(x)| dx \Big) \\
&= Ks^{\frac{1}{2}+\alpha}2^\alpha \Big(\int_{-\infty}^{\infty} |x|^\alpha |\psi(x)| dx + \Big(\frac{|u-v|}{s}\Big)^\alpha \int_{-\infty}^{\infty} |\psi(x)| dx \Big) \\
&= As^{\frac{1}{2}+\alpha}\Big(1 + \Big(\frac{|u-v|}{s}\Big)^\alpha\Big).
\end{aligned}$$

The second equality follows by a change of variable $x \cdot s = t - u$, and the second inequality follows by the inequality $|a+b|^n \leq 2^n(|a|^n + |b|^n)$. If f is uniformly Lipschitz-α over $]a,b[$ this equation holds for all $\nu \in]a,b[$ and particularly for $\nu = u$.

Theorem 2.4.2 ([1]) *If $f \in L^2(\mathbb{R})$ is locally Lipschitz-α at ν and $\psi(t)$ has $N > \alpha$ vanishing moments, then there exists an constant $A > 0$ such that;*

$$|Wf(u,s)| \leq As^{\alpha+\frac{1}{2}}\Big(1 + \big(\frac{|u-v|}{s}\big)^{\alpha}\Big). \tag{2.4.2}$$

If f is uniformly Lipschitz-α in $]a,b[$ and $\psi(t)$ has $N > \alpha$ vanishing moments, then there exists an constant $A > 0$ such that;

$$|Wf(u,s)| \leq As^{\alpha+\frac{1}{2}}. \tag{2.4.3}$$

□

Even if it is a nice result, it will be of little importance in analyzing the regularity of a function. The Lipschitz-regularity is a priori unknown. The next theorem will enable to estimate the Lipschitz-regularity.

Theorem 2.4.3 ([1]) *Assume $\psi(t)$ has n vanishing moments, and $f \in L^2[a,b]$ satisfies*

$$|Wf(u,s)| \leq As^{\alpha+\frac{1}{2}}\Big(1 + \big(\frac{|u-v|}{s}\big)^{\alpha'}\Big) \tag{2.4.4}$$

for $\alpha' < \alpha < n$ and $\alpha \notin \mathbb{N}$. Then f is Lipschitz-α at v.
Assume that f is bounded and $Wf(u,s)$ satisfies Eq(2.4.3) for an $\alpha < n$, $\alpha \notin \mathbb{N}$. Then f is uniform Lipschitz-α in $[a+\epsilon, b-\epsilon]$.

□

This result will be the main tool in analyzing the Lipschitz-regularity of a function. The α can be computed by the value of the wavelet-transform at two different scales. For $s_0 < s_1$, α can be computed by the equation;

$$\alpha = \frac{\log_2(|Wf(u,s_1)|) - \log_2(|Wf(u,s_0)|)}{\log_2(s_1) - \log_2(s_0)} - \frac{1}{2}, \tag{2.4.5}$$

in the uniform case. In Sect.(2.5) it will be verified that in certain cases a similar equation may be used to estimate the local Lipschitz-regularity of a function.

Example 2.4.1 *In Sect.(2.2) the wavelet-transform of the Heaviside was computed. For $s_0 < s_1$ the global Lipschitz exponent around 0 is;*

$$\begin{aligned}\alpha + \frac{1}{2} &= \frac{\log_2(|Wf(0,s_1)|) - \log_2(|Wf(0,s_0)|)}{\log_2(s_1) - \log_2(s_0)} \\ &= \frac{\log_2(\sqrt{s_1}) - \log_2(\sqrt{s_0})}{\log_2(s_1) - \log_2(s_0)} = \frac{1}{2} \\ &\Rightarrow \alpha = 0. \end{aligned} \tag{2.4.6}$$

This agrees with the result found in Ex.(2.2.1).

□

2.5 Wavelet-Transform Modulus-Maximum.

In this section some useful properties of the *wavelet-transform modulus-maximum* will be discussed, and the concepts *maxima-line* and *maxima-tree* introduced.

In the previous section the relation between the Lipschitz-regularity and the CWT was established. In order to find the Lipschitz-regularity of a function, the previous section verified that this could be estimated by computing the decay of the wavelet-transform at all points. This section will prove that it is sufficient to compute the wavelet-transform at the points of worst behaviour; at the modulus-maximum of the wavelet-transform. This will introduce a new problem of connecting modulus-maxima across scales. The maxima-tree will create a hierarchal structure which relate modulus-maximum across scales.

This section will prove that irregular structures in a signal can be found among the modulus-maximum of the CWT. By the help of an example it will be illustrated that all wavelets are not appropriate to use for across-scale analysis, i.e. to compare modulus-maximum of the wavelet-transform at different scales. It will be verified that all Gaussian derived wavelets are suitable for across-scale analysis.

The section begins with a theorem implying that irregularities can be found among the modulus-maximum of the CWT. The concept of *maxima-line* and *maxima-tree* is defined and discussed.

The first theorem implies that it suffices to search for irregularities among the modulus-maximum of the CWT.

Theorem 2.5.1 ([1]) *Let $\psi \in C^n(\mathbb{R})$ have fast decay and n vanishing moments, and $\psi(t)$ be compactly supported. Let $f \in L^1[a,b]$. If there exist an $s_0 > 0$ such that $\forall s < s_0$ $|Wf(u,s)|$ has no local maxima for $u \in [a,b]$, then f is uniformly Lipschitz-n on $[a+\epsilon, b-\epsilon]$, for an $\epsilon > 0$.*

□

This theorem implies that all irregularities can be found among the modulus-maximum of the wavelet-transform at fine scales. It does not imply that all local maxima at fine scales correspond to an irregularity. In order to find which of the local-maxima correspond to an edge, one has to estimate the Lipschitz-regularity of the structure in the function which causes the modulus-maximum. From the discussion in the previous section it follows that in order to estimate the Lipschitz-regularity, one has to know the value of the modulus-maximum at least at two different scales. One has to relate modulus-maximum occurring from similar irregularities in $f(t)$, i.e. find the *maxima-line* converging to a modulus-maximum.

Definition 2.5.1 (Maxima Line & Maxima-tree.) *A maxima-line is any connected curve in the time-scale space $\mathbb{R}\times\mathbb{R}_+$ along all points are a modulus-maximum of the wavelet-transform. The term maxima-tree will be used to represent the set of all maxima-lines in a signal.*

□

Finding maxima-lines and maxima-trees will be an important issue in this thesis. In fact, the maxima-tree may be used as the basis for a simple edge-detector. This will be discussed in Ch.(5). The next example illustrates the maxima-line of the wavelet-transform for two different wavelets, illustrating an important difference between wavelets.

Example 2.5.1 *Consider the wavelet in Fig.(2.8d). The wavelet is a spline-wavelet of order 1, and given by;*

$$\psi(t) = \begin{cases} 0 & \text{for } |t| \geq 1 \\ -(t+1) & \text{for } t \in]-1,-1/2[\\ t & \text{for } t \in [-1/2, 1/2] \\ -(t-1) & \text{for } t \in]1/2, 1[\end{cases}.$$

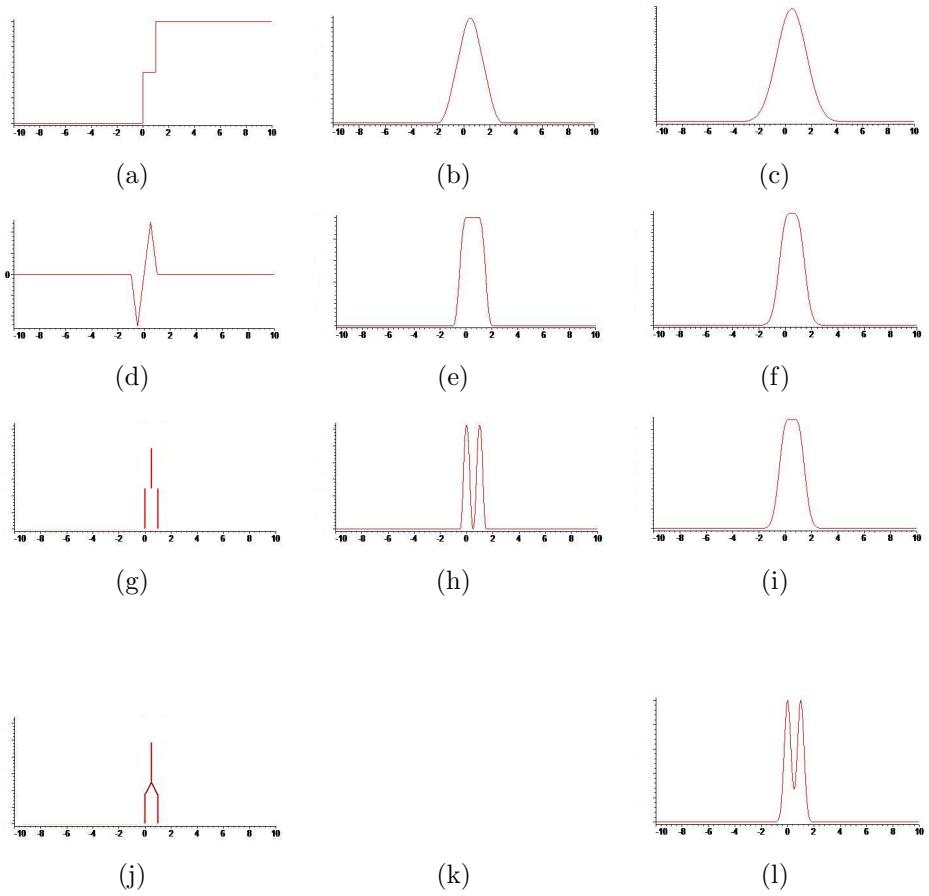


Figure 2.8: (a) The signal $f(t)$. (b)-(e)-(h) The CWT of $f(t)$ with respect to the spline-wavelet. The CWT is computed for resp. $s > 1$, $s = 1$ and $s < 1$. (c) - (f) - (i) The CWT of $f(t)$ with respect to the the wavelet corresponding to the 1.st derivative of the Gaussian. The CWT is computed for resp. $s > 1$, $s = 1$ and $s < 1$. (g) - (j) The maxima-line corresponding to resp. spline-wavelet and Gaussian wavelet.

Assume the signal is given by $f(t) = u(t)+u(t-1)$, *see Fig.(2.8a). For* $s > 1$ *the CWT of both wavelets treat the two step-edges as one edge. Since both edges are of the same strength the position of the modulus-maximum is at* $t_0 = 1/2$. *This is illustrated in Fig.(2.8(b),(c)). For* $s = 1$ *the CWT differs for the two wavelets. By definition of modulus-maximum the local maxima*

should be strict from either left or right. It can be verified that the value of the CWT at $s = 1$ *equal* $1/4$ *for all* $u \in [0, 1]$. *The modulus-maximum of the CWT with respect to the spline-wavelet is therefore at* 0 *and* 1. *The modulus-maximum of CWT with respect to the Gaussian wavelet is at* $1/2$. *For* $s < 1$ *the modulus-maximum for the CWT with respect to spline-wavelet is at* 0 *and* 1. *For the CWT with respect to the Gaussian wavelet, the modulus-maximums split and converge towards* 0 *and* 1 *as illustrated in Fig.(2.8i,l). Fig.(2.8,g,j) display the maxima-line for the two CWT. Note that one of them is disconnected at* $s = 1$ *while the other is connected for all* s. *I.e. the maxima-line for the Gaussian-derived wavelet converges towards fine scales, while for the spline wavelet the maxima-line does not necessarily converge. The statements of the behaviour of the maxima-line with respect to the will be verified by Th.(2.5.2) and in Sect.(5.1).*

□

This example illustrates an important difference between wavelets. In order to estimate the Lipschitz-regularity by the CWT one need to relate modulus-maximum at different scales. If the maxima-lines are disconnected it makes it difficult to relate modulus-maximum across scales which should be used to estimate the Lipschitz-regularity. In addition the length of the maxima-line converging to a modulus-maximum will be an important characteristic for noisy signals. This will be discussed in Sect.(5.1). The following theorem guarantees that the maxima-line is connected for the Gaussian-derived wavelets.

Theorem 2.5.2 ([1]) *Assume* $\psi(t) = (-1)^n \theta^{(n)}(t)$ *where* $\theta(t)$ *is the Gaussian-function. For any* $f \in L^2(\mathbb{R})$, *the modulus-maximum of* $Wf(u, s)$ *belong to connected maxima-line that are never interrupted as the scale decreases.*

□

Proposisjon 2.5.1 *Assume* $f \in L^2(\mathbb{R})$, *and* $\psi(t) = -\theta'(t)$. *Then the function;*

$$Wf(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$$

is continuous at $\mathbb{R} \times \mathbb{R}_+$.

Proof: Note that the function $\psi_s(t) : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is continuous. For an arbitrary point $(u_0, s_0) \in \mathbb{R} \times \mathbb{R}_+$ it follows by Schwarz-inequality that for all

$(u,s) \in \mathbb{R} \times \mathbb{R}_+$;

$$|Wf(u,s) - Wf(u_0,s_0)| \leq \| f \| \| \psi_s(t-u) - \psi_{s_0}(t-u_0) \| .$$

Since the norm is continuous, it follows that $Wf(u,s)$ is continuous at $\mathbb{R}\times\mathbb{R}_+$.

□

By the theorem and the property one has the following important consequences with respect to the Gaussian derived wavelets;

- All maxima-lines converges towards finer scale.
- All maxima-lines converges to at least one modulus-maximum.
- Along a maxima-line small changes of scale will results in a small change of the value of the modulus-maximum.

These properties will be important in Ch.(5).

Two questions come forward at this point, which wavelets should be used and how one may estimate the Lipschitz-regularity at a point using Eq.(2.4.4). Th.(2.5.1) verified that the more vanishing moments of a wavelet, the larger number of different kind of edges can be detected. On the other hand, the more vanishing moments, the larger number of modulus-maximum. E.g. the response of a step-edge to the Mexican-Hat had two modulus-maximum, etc. To reduce the number of necessary computations it is therefore necessary to keep as few vanishing moments as possible. In this thesis, step-edges are typically the interesting edges in the signal. The wavelet corresponding to the 1.st derivative of the Gaussian will therefore be the preferred wavelet.

The obstacle with using Eq.(2.4.4) to estimate the pointwise Lipschitz-regularity, is that the estimate requires both α and α'. To circumvent this problem, one may use the *Cone of Influence*. Assume $supp(\psi) = [-C, C]$, for some $C \in \mathbb{R}_+$. Then $supp(\psi_{u,s}) = [u - Cs, u + Cs]$. The Cone of Influence of a point ν is defined as;

$$|u - \nu| \leq Cs.$$

Assume ν is the point at which a maxima-line converge as $s \to 0$. If all elements u of the maxima-line converging towards ν are inside the Cone of Influence of ν, then one may reduce Eq.(2.4.4) to the following simple form;

$$|Wf(u,s)| \leq A' s^{\alpha+1/2}.$$

I.e. if the spatial position of the elements of a maxima-line does not changes considerably with respect to the scale, one can use Eq.(2.4.5) to estimate the Lipschitz-regularity at ν.

2.6 Error-Analysis.

In this section errors one may encounter in numerical analysis will be discussed, in particular errors which may influence the estimate of the Lipschitz-regularity.

Theoretically one may estimate the Lipschitz-regularity for all $s > 0$. This can typically not be done in practice, nor does it make much sense. The signal itself is a piecewise constant approximation of "something". I.e. the signal is Lipschitz-0 between most sample-points. Lower bound for s is needed in order to obtain the wanted information in the signal. Computing the exact value of the wavelet-transform will typically be very time-consuming. To reduce computational efforts, an approximation of the wavelet-transform will be introduced. This approximation will be pointless at small scales. These lower bounds will cause problems for estimating the Lipschitz-regularity. Recall that the Lipschitz-regularity was estimated by the CWT as $s \to 0$. The lower bounds for the scale introduces errors in the estimate of α. Some errors are discussed and estimated.

First the error caused by approximation of the CWT, oscillations and smoothening of edges is discussed. Lower bound for the scale is found. At the end it will be discussed the effect of multiple singularities in the signal, resulting in an upper bound for the scale. The estimates for the lower and upper bound will be used in an example in Sect.(2.7).

2.6.1 Approximation errors.

This section presents a formula which approximates the continuous wavelet-transform, and discusses errors occurring as a consequence of this approximation.

Computing the exact value of the wavelet-transform is typically an expensive operation and is typically not obtainable in practice. This motivates the introduction of an approximation of the wavelet-transform. The pay-off for using such approximations is that it does not make sense to compute the wavelet-transform at small scales.

The presentation begins with deriving an approximation-formula for $Wf(u,s)$ based on the Trapezoid-formula. With the help of an example it will be discussed why this approximation is senseless at small scales.

Assume $f(t) = 0$ for $t \notin]a,b[$ with $-\infty < a < b < \infty$. Assume for simplicity that $a = 0$ and $b = N$ for some integer N. The Trapezoid-approximation of the function $f(t)$ with uniform spacing and sampling-distance 1 is [**?**];

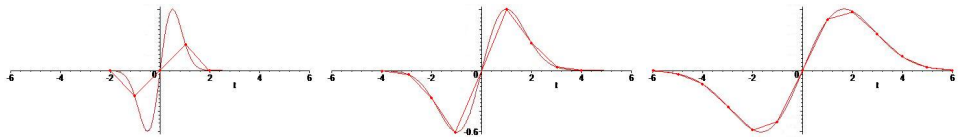
$$T(f) = \frac{1}{2}\sum_{n=0}^{N-1}\Big(y_n + y_{n+1}\Big) = \frac{y_0 + y_N}{2} + \sum_{n=1}^{N-1} y_n.$$

Since $f(0) = f(N) = 0$ the corresponding estimate for the wavelet-transform at integer-points $u_n \in \{u_0, ..., u_N\}$;

$$Wf(u_n, s) = \sum_{n=0}^{N-1} f(n)\frac{1}{\sqrt{s}}\psi\Big(\frac{n-u_n}{s}\Big) = f * \tilde{\psi}_s(u_n),$$

where $*$ denoted the discrete convolution, and $\tilde{f}(t) = f(-t)$. Calculating the value of the wavelet-transform by means of discrete convolution is particular nice for implementation in MatLab. There exist several convolution-commands in MatLab which is fast, e.g. "conv2".

This approximation will cause havoc for the decay of the wavelet-transform at small scales. It can be verified that with resp. to a step-edge, then for $s > 6/4$ the decay of $log_2(|Wf(u,s)|)$ is approximately $1/2$, while for $s < 6/4$ it starts to decay (a lot) faster that $1/2$. One factor which causes havoc at small scales is the form the wavelet takes at small scales. Consider Fig.(2.9)



(a) (b) (c)

Figure 2.9: Wavelets corresponding to the 1.st derivative of the Gaussian at three different scales. The smooth line corresponds to the actual line, while the line-segments are what taken into account by the approximation.

illustrating the wavelets corresponding to the 1.st derivative for three different scales $(a)s = 1/2, (b)s = 1, (c)s = 6/4$, and the form of the wavelets

taken into account by the Trapezoid-formula with integer-approximation. At large scales the approximated wavelet and the wavelet is only slightly different. As the scale decrease the difference between the actual wavelet and the "wavelet" used by the approximation increases. This difference causes the sudden decay of the CWT at small scales. In this thesis the lower bound for the scale will be assumed to be $s \geq 6/4$. Because this lower bound will typically be dominated by other lower bounds discussed in the next sections, no deep analysis of this error is done.

If one has to analyze the signal for small scales, this problem may be solved by using a shorter interval in the Trapezoid-approximation of CWT.

2.6.2 Oscillating signals.

This section will present the effect of oscillations in a signal.

Oscillations in signals represent another difference between theory and numerical analysis. In theory one can avoid the problem of oscillations in a signal by using sufficiently small scales. As discussed, one does not have the possibility in the numeric case. Oscillations are present in signals and are typically caused by texture and noise. Sect.(4.1) will discuss *additive Gaussian White Noise* in signals.

The discussion of oscillations is done by considering the oscillations as being trigonometrically distributed, i.e. can be expressed as a sum of cosines. This can not be used to represent noise and texture which is randomly distributed, but will be used as a model of these.

The relation between cosine-distributed oscillations and the CWT is investigated, and estimates are found. These estimates can be used to find lower bound for the scale such that the oscillations are not "felt" by CWT.

Fig.(2.10) display the decay of the wavelet-transform at a modulus-maximum of a step-edge and a step-edge with added oscillations given as $1/6\cos(4\pi t) + 1/5\sin(\pi t)$. This demonstrates that oscillations affect the decay of the wavelet-transform. One may observe from Fig.(2.10) that this effect is mostly present for small scales.

The general expression for oscillations will be assumed to be a sum of

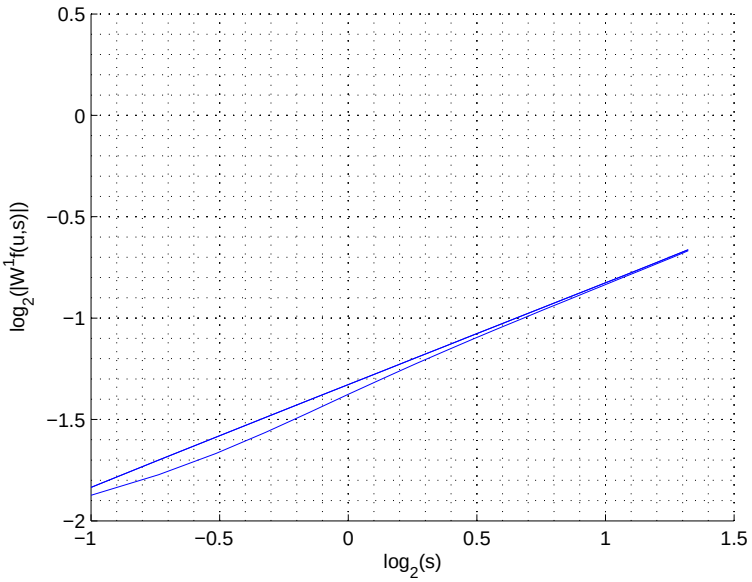


Figure 2.10: The decay of the wavelet-transform a step-edge with and without added oscillations.

trigonometric functions, i.e.

$$N(t) = \sum_{i=1}^{k} C_{\gamma_i,\omega_i}(t) = \sum_{i=1}^{k} \chi_{[a_i,b_i]}\gamma_i \cos(\omega_i t) \tag{2.6.1}$$

$$\text{with } \cos(\omega_i a_i) = \cos(\omega_i b_i) = 0 \text{ and } -\infty < a_i < b_i < \infty. \tag{2.6.2}$$

The latter requirement is included to avoid artificial step-edges. The aim is to find an estimate on the error caused by oscillations in a function, and from this derive a lower bound for the scale which make the influence of oscillations small.

Statement 2.6.1 *Let* $f \in L^2(\mathbb{R}), \psi(t) = -\frac{d\theta(t)}{dt}$, $\theta(t)$ *is the Gaussian, and* $N(t) = \sum_{i=1}^{k} C_{\gamma_i,\omega_i} \in L^2(\mathbb{R})$ *be defined as in Eq.(2.6.1). Then;*

$$|W^{(1)}f_0(u,s) - W^{(1)}(f_0+N)(u,s)| = |W^{(1)}N(u,s)| = |\sum_{i=1}^{k} W^{(1)}C_{(\gamma_i,\omega_i)}(u,s)|,$$

and

$$|W^{(1)}C_{\gamma,\omega}(u,s)| \leq \gamma\omega s^2 s^{-\frac{1}{2}} e^{-\frac{\omega^2 s^2}{2}}. \tag{2.6.3}$$

Proof:

$$\begin{aligned}
|W^{(1)}C_{\gamma,\omega}(u,s)| &= |\frac{\gamma}{\sqrt{s}}\int_{\mathbb{R}} \chi_{[t_1,t_2]}\cos(\omega t)\psi(\frac{t-u}{s})dt| \leq |\frac{\gamma}{\sqrt{s}}\int_{\mathbb{R}} \cos(\omega t)\psi(\frac{t-u}{s})dt| \\
\boxed{\tau = t-u} &= \frac{\gamma}{\sqrt{s}}|\int_{\mathbb{R}} \cos(\omega\tau + u\omega)\psi(\frac{\tau}{s})d\tau| \\
&= \frac{\gamma}{\sqrt{s}}|\cos(u\omega)\int_{\mathbb{R}} \cos(\omega\tau)\psi(\frac{\tau}{s})d\tau - \sin(u\omega)\int_{\mathbb{R}} \sin(\omega\tau)\psi(\frac{\tau}{s})d\tau|
\end{aligned}$$

Since $\cos(\omega\tau)\theta'(\frac{\tau}{s})$ is an odd function the first integral equals 0, so

$$\begin{aligned}
|W^{(1)}C_{\gamma,\omega}(u,s)| &= \frac{\gamma}{\sqrt{s}}|\sin(u\omega)\int_{\mathbb{R}} \sin(\omega\tau)\psi(\frac{\tau}{s})d\tau| \\
&= \frac{s\gamma}{\sqrt{s}}|\sin(u\omega)\big[sin(\omega\tau)\theta(\frac{\tau}{s})\big]_{\tau=-\infty}^{\infty} - \omega\sin(u\omega)\int_{\mathbb{R}} \cos(\omega\tau)\theta(\frac{\tau}{s})d\tau| \\
&= \frac{s\gamma}{\sqrt{s}}|\omega\sin(u\omega)\int_{\mathbb{R}} \cos(\omega\tau)\theta(\frac{\tau}{s})d\tau| \leq \frac{s\gamma\omega}{\sqrt{s2\pi}}|\int_{\mathbb{R}} \cos(\omega\tau)e^{-\frac{\tau^2}{2s^2}}d\tau| \\
&= \frac{2s\gamma\omega}{\sqrt{s2\pi}}|\int_0^{\infty} \cos(\omega\tau)e^{-\frac{\tau^2}{2s^2}}d\tau| = \frac{2s\gamma\omega}{\sqrt{s2\pi}}\frac{1}{2}\sqrt{2s^2\pi}e^{-\frac{\omega^2 s^2}{2}} \\
&= \frac{s^2\gamma\omega}{\sqrt{s}}e^{-\frac{\omega^2 s^2}{2}}
\end{aligned}$$

Where the solution of the integral follows from cosine Fourier-transform.

□

This statement verifies that the error caused by oscillations is closely related to the amplitude and frequency of the oscillations. I.e. high frequencies and low amplitude yields little error, while low frequencies and high amplitude yields a large error. To finish this topic, a similar result will be derived for arbitrary Gaussian derivatives.

Statement 2.6.2 *Let $f \in L^2(\mathbb{R})$ and $N(t)$ be defined as in Statement 2.6.1, and let $\psi(t) = (-1)^n \frac{d\theta}{dt^n}$, where $\theta(t)$ is that Gaussian. Then*

$$|W^{(n)}C_{\gamma,\omega}(u,s)| = (s\omega)^{n-1}|W^{(1)}C_{\gamma,\omega}(u,s)| \leq (s\omega)^{n-1}\gamma\omega s^2\frac{1}{\sqrt{s}}e^{-\frac{\omega^2 s^2}{2}} \tag{2.6.4}$$

Proof: This will be proved by induction, where the induction start is done in Stat.(2.6.1). Assume that Eq.(2.6.4) holds $\forall k < n$. Then;

$$|W^{(n)}C_{\gamma,\omega}(u,s)| = \frac{\gamma}{\sqrt{s}}|\int_{\mathbb{R}} \cos(\omega t)\psi(\frac{t-u}{s})dt| = \frac{\gamma}{\sqrt{s}}|\int_{\mathbb{R}} \cos(\omega\tau + u\omega)\psi(\frac{\tau}{s})dt|$$
$$= \frac{\gamma}{\sqrt{s}}|\cos(u\omega)\int_{\mathbb{R}} \cos(\omega\tau)\psi(\frac{\tau}{s})d\tau - \sin(u\omega)\int_{\mathbb{R}} \sin(\omega\tau)\psi(\frac{\tau}{s})d\tau|.$$

Assume that $n = 2m$ for some $m \in \mathbb{N}$. Then $\cos(\omega\tau)\theta^{(2m)}(\tau/s)$ is even and $\sin(\omega\tau)\theta^{(2m)}(\frac{\tau}{s})$ is odd and therefore the second integral equals 0. So

$$|W^{(n)}C(u,s)| = \frac{\gamma}{\sqrt{s}}|\cos(u\omega)\int_{\mathbb{R}} \cos(\omega\tau)\theta^{(2m)}(\frac{\tau}{s})d\tau|$$
$$= |\frac{s\gamma}{\sqrt{s}}\big[\cos(\omega\tau)\theta^{(2m-1)}(\frac{\tau}{s})\big]_{\tau=\infty}^{\infty} + \frac{s\omega\gamma}{\sqrt{s}}\int_{\mathbb{R}} \sin(\omega\tau)\theta^{(2m-1)}(\frac{\tau}{s})d\tau|$$
$$= s\omega(\frac{\gamma}{\sqrt{s}}\int_{\mathbb{R}} \sin(\omega\tau)\theta^{(2m-1)}(\frac{\tau}{s})d\tau) = s\omega W^{(2m-1)}C_{\gamma,\omega}(u,s).$$

Assume that $n = 2m+1$. Then $\cos(\omega\tau)\theta^{(2m+1)}(\tau/s)$ is odd and $\sin(\omega\tau)\theta^{(2m+1)}(\frac{\tau}{s})$ is even and therefore the first integral equals 0. So

$$|W^{(n)}C(u,s)| = \frac{\gamma}{\sqrt{s}}|\sin(u\omega)\int_{\mathbb{R}} \sin(\omega\tau)\theta^{(2m+1)}(\frac{\tau}{s})d\tau|$$
$$= |\frac{s\gamma}{\sqrt{s}}\big[\sin(\omega\tau)\theta^{(2m)}(\frac{\tau}{s})\big]_{\tau=\infty}^{\infty} + \frac{s\omega\gamma}{\sqrt{s}}\int_{\mathbb{R}} \cos(\omega\tau)\theta^{(2m)}(\frac{\tau}{s})d\tau|$$
$$= s\omega(\frac{\gamma}{\sqrt{s}}\int_{\mathbb{R}} \cos(\omega\tau)\theta^{(2m)}(\frac{\tau}{s})d\tau) = s\omega W^{(2m)}C_{\gamma,\omega}(u,s).$$

□

The expressions for the influence of oscillations with respect to the CWT serve two purposes. First, the error caused by oscillations may be chosen arbitrary small by choosing a high lower bound for the scale. Secondly the statements imply that the error is closely related to the frequency and amplitude of the distributions.

2.6.3 Smoothed signals.

There is one typical difference between (applied) signals and (theoretic) functions. Theoretic edges such as step-edge and angle-edge would typically not

occur in practice. Physical phenomenon can not include instant changes of state, but rather a "smooth" transition between states. Still we would like to treat them both as a step-edge, i.e. both should have Lipschitz-exponent 0. The following theorem estimates the error caused by smoothening of an edge.

Theorem 2.6.1 ([1]) *Let $\psi(t) = (-1)^n \theta^{(n)}(t)$ with θ the Gaussian. If $f(t) = f_0 * g_\sigma$ with f_0 uniformly Lipschitz $\alpha < n$ on $[v-h, v+h]$, and g_σ given by*

$$g_\sigma(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}}.$$

Then there exists an $A \in \mathbb{R}$ such that

$$\forall (u,s) \in [v-h, v+h] \times \mathbb{R}_+ \ , \ |Wf(u,s)| \leq A s^{\alpha+1/2} \Big(1 + \frac{\sigma^2}{s^2}\Big)^{-(n-\alpha)/2}. \quad (2.6.5)$$

□

One may observe from the Theorem that pre-knowledge of α is required in order to use Eq.(2.6.5). This is generally unknown. The interesting observation which may be done from Eq.(2.6.5) is that if $\frac{\sigma^2}{s^2}$ is small, the calculated α-value is small. I.e. at large scales the error caused by smoothening is small. This reflects the zooming-property of the wavelet-transform. If one observe a smoothened edge from a distance is appears to be an instant-edge. Another useful observation is that for a step-edge and the wavelet $\psi^1(t)$, the theorem gives that the wavelet-transform of a smoothened step-edge decays faster than a step-edge as the scale decreases.

2.6.4 Influence of multiple edges.

The final part of the error-analysis is to consider the effect of multiple singularities. The previously discussed causes of error have resulted in lower bounds for the scale. For multiple singularities an upper bound for the scale is achieved.

Consider Fig.(2.11). The Lipschitz-regularity at $\nu = 5$ and $\nu = 10$ is 0. The signal in Fig.(2.11) is used as an example of the influence of multiple edges of the decay of the wavelet-transform. The upper line in Fig.(2.12) display the decay of the wavelet-transform at $\nu = 5$. The lower display a line

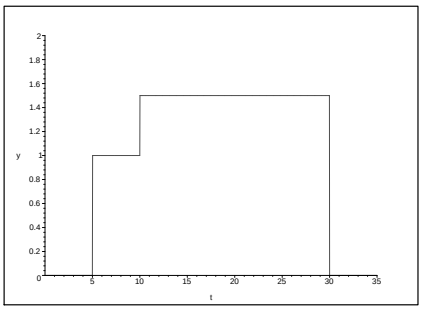


Figure 2.11: Piecewise constant signal with a jump at $\nu = 5$ and $\nu = 10$. The Lipschitz-regularity at both steps equal 0.

decaying as 1/2 corresponding to Lipschitz-exponent 0. From Fig.(2.12) one may observe that the decay of the wavelet-transform at $\nu = 5$ at the small scales corresponds to a step-edge. As the scale increases above $log_2(s) \approx 0.5$

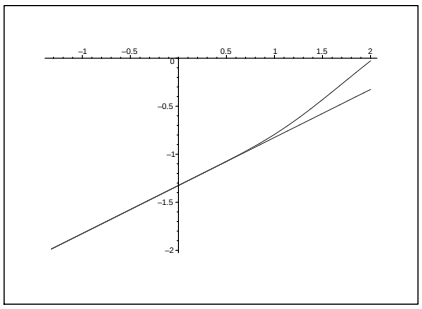


Figure 2.12: The upper line corresponds to decay of the wavelet-transform of the signal in Fig.(2.11) at $\nu = 5$. The lower line correspond to an edge which is Lipschitz-0 at all scales.

the the decay of the wavelet-transform differs. Assume

$$f_0(t) = \begin{cases} 0 & \text{for } t \notin [0, x_2] \\ a & \text{for } t \in [0, x_2] \end{cases}$$

and

$$f_1(t) = \begin{cases} 0 & \text{for } t \notin [0, x_2] \\ a & \text{for } t \in [0, x_1] \\ b & \text{for } t \in [x_1, x_2] \end{cases}$$

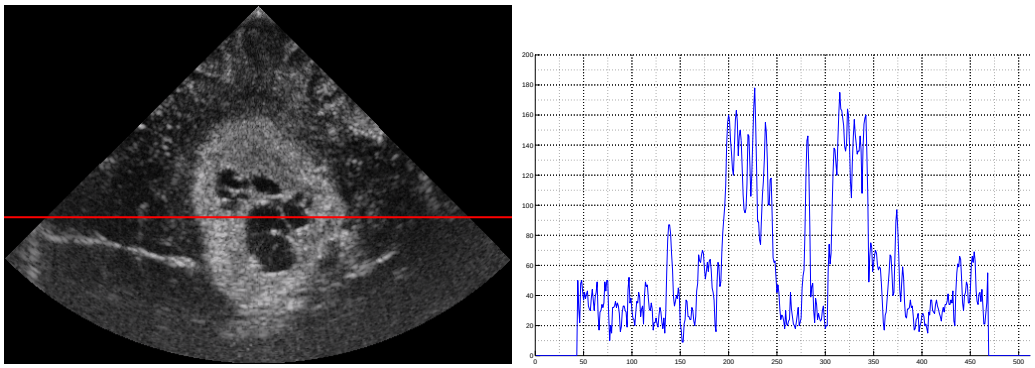
where $0 < x_1 < x_2$. Then the difference of the wavelet-transform at 0 of the two signals is

$$\begin{aligned} |Wf_0(0,s) - Wf(0,s)| &= |\tfrac{1}{\sqrt{s}} \int_{\mathbb{R}} f_0(t)\psi(\tfrac{t}{s})dt - \tfrac{1}{\sqrt{s}} \int_{\mathbb{R}} f_1(t)\psi(\tfrac{t}{s})dt| \\ &= |\tfrac{a}{\sqrt{s}} \int_0^{x_2} \psi(\tfrac{t}{s})dt - \tfrac{a}{\sqrt{s}} \int_0^{x_1} \psi(\tfrac{t}{s})dt - \tfrac{b}{\sqrt{s}} \int_{x_1}^{x_2} \psi(\tfrac{t}{s})dt| \\ &= |\tfrac{a-b}{\sqrt{s}} \int_{x_1}^{x_2} \psi(\tfrac{t}{s})dt| \end{aligned}$$

It follows that if the support of the wavelet includes the interval $[x_1,x_2]$, i.e. $supp(\psi_s(t)) \cap [x_1, x_2] \neq \emptyset$, it affects the wavelet-transform. So for a signal with multiple edges, one has to ensure that the scale is low enough such that the support of the wavelet does not include the other edges.

2.7 Example of estimating the Lipschitz-regularity.

To finish off this chapter, the estimates developed in the previous section will be applied for a 1-D ray in the ultrasound image, see Fig.(2.13) First, the



(a) (b)

Figure 2.13: The image in (a) is the ultrasound-image which is investigated in this thesis. The horizontal line in (a) correspond to the signal in (b).

edges to be analysed has to be found. The edges which will be analysed are the modulus-maximum at $u \in \{43, 133, 147, 167, 195, 247, 275, 288, 307, 345, 381, 467\}$ in Fig.(2.14). The target is to see which of these edges correspond to a step-edge, and if one can imrove the estimate of the Lipschitz-regularity by using the error-sources discussed in the previous section. Fig.(2.14) is the wavelet-transform of the signal in Fig.(2.13) at scale $s = 6$. The next step is to find

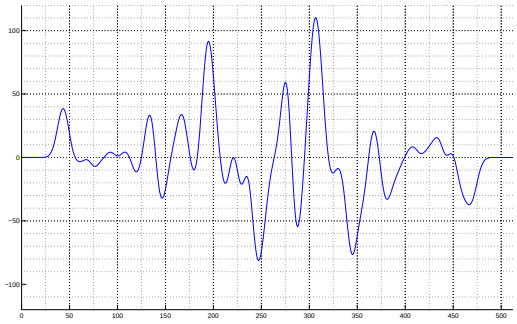


Figure 2.14: The wavelet-transform for the signal in Fig.(2.13b) at scale $s = 6$. The modulus-maximum above 25 will be the edges to be analysed.

the maxima-lines corresponding to these modulus-maximum, and find which points the modulus-maximum above converge to at small scales. In Fig.(2.15) the maxima-lines for the scales $s \in [1, 6]$ are plotted, and these converges to the points $\{43, 136, 142, 167, 197, 245, 279, 284, 307, 344, 376, 469\}$. First one

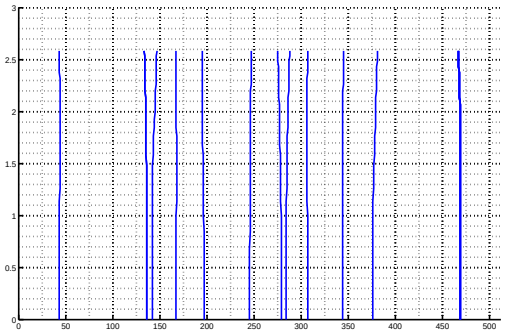


Figure 2.15: The maxima-lines for $s \in [1, 6]$ for the signal in Fig.(2.13b).

has to find an estimate for the oscillations in the figure. Since the oscillations are unknown, the writer choose to approximate the oscillations on the form as in Eq.(2.6.1), where γ_i is chosen as the average amplitude, and ω_i are the average frequency. In the signal in Fig.(2.13b), it seems as the oscillations differs at high values and low values for the signal. The writer assume that the oscillations in $U = [43, 192] \cup [253, 275] \cup [345, 469]$ is equally distributed,

and the oscillations in $V = [193, 252] \cup [276, 244]$ is equally distributed. Further the writer assumes that the oscillations in both domains are distributed as a single term in Eq.(2.6.1). By averaging the amplitude and the frequency in the two domains, assume that

$$N(t) = \chi_U 8.5 \cos(1.1t) + \chi_V 25.33 \cos(0.7t)$$

These values are used to find the lower bounds due to oscillations. One has to compute three estimates. The edges in U sufficiently far from V are only affected by $\chi_U 8.5 \cos(1.1t)$ and similarly for the edges in V far from U. The edges lying at or close the transition between U and V are affected by $\chi_U 8.5 \cos(1.1t) + \chi_V 25.33 \cos(0.7t)$. The lower bounds found for the wavelet corresponding to the 1.st derivative of the Gaussian are;

$$\begin{aligned} s_0 &> 3.825 \quad \text{for the edges in } U \\ s_0 &> 6.478 \quad \text{for the edges in } V \\ s_0 &> 6.478 \quad \text{for the edges in close to } U \text{ and } V. \end{aligned}$$

The corresponding lower bounds for Mexican Hat are, according to Thm.(2.6.2);

$$\begin{aligned} s_0 &> 4.489 \quad \text{for the edges in } U \\ s_0 &> 7.622 \quad \text{for the edges in } V \\ s_0 &> 7.622 \quad \text{for the edges in close to } U \text{ and } V. \end{aligned}$$

The last two lower bounds equals for at least the 6 first digits in both expressions. The upper bound, should according to Sect.(2.6), be the distance from the edge to of its closest neighbour-edge. Table(2.1) display the upper and lower bounds for all the edges which are analysed. In Fig.(2.16) the decay of the wavelet-transform is plotted for the different edges. Within their respective allowed values for the scales, shown in Fig.(2.1) the decays are shown in Table(2.2)

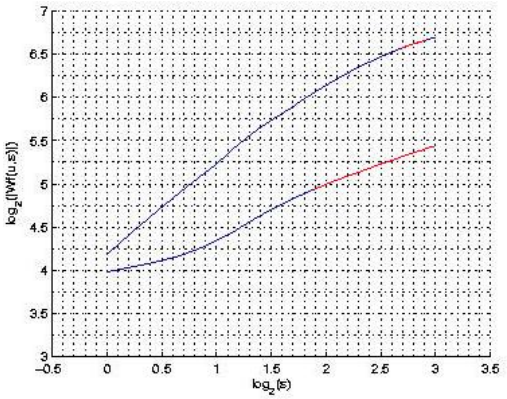
This example illustrates the difficulties in estimating the Lipschitz-regularity of a modulus-maximum. One may observe that even with error-estimates it may be difficult to get good estimates for the Lipschitz-regularity, and that the estimate is worse without esimating which scales should be taken into account. This uncertainty will cause problems for edge-detectors taking advantage of the Lipschitz-regularity. This will be discussed in Ch.(6) and Ch.(7).

ν	s_0	s_1	s_2	$[log_2 s_0, log_2 s_2]$	$[log_2 s_1, log_2 s_2]$
43	3.825	4.489	23.25	[1.935,4.539]	[2.166,4.539]
136	3.825	4.489	1.5	$\emptyset$	$\emptyset$
142	3.825	4.489	1.5	$\emptyset$	$\emptyset$
167	3.825	4.489	6.25	[1.935,2.644]	[2.166,2.644]
197	6.478	7.622	7.5	[2.696,2.907]	$\emptyset$
245	6.478	7.622	8.5	[2.696,3.087]	[2.930,3.087]
279	3.825	4.489	1.25	$\emptyset$	$\emptyset$
284	3.825	4.489	1.25	$\emptyset$	$\emptyset$
307	6.478	7.622	5.75	$\emptyset$	$\emptyset$
344	6.478	7.622	8	[2.696,3]	[2.930,3]
376	3.825	4.489	8	[1.935,3]	[2.166,3]
469	3.825	4.489	23.25	[1.935,4.539]	[2.166,4.539]

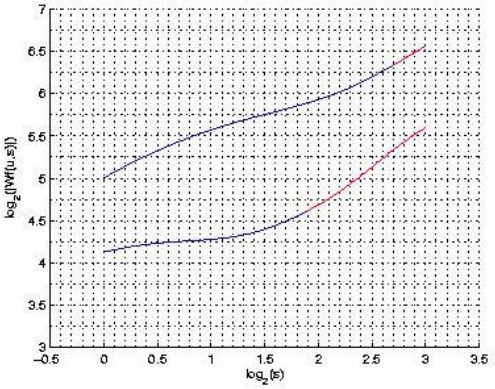
Table 2.1: Table with the bounds for the scales used to estimate the Lipschitz-regularity of the signal $f(t)$ if Fig.(2.13b). s_0 is the bounds due to $-\theta'(t)$, s_1 is bounds due to $\theta^{(2)}(t)$, and s_2 is the upper bound due to other singularities.

ν	Δ	Δ'	$\Delta^{(2)}$
43	0.44	0.64	
167	0.35	0.45	
197	0.41	0.82	
245	0.6	0.99	
344	0.75	0.67	1.48
376	0.461	0.53	
469	0.91	0.91	

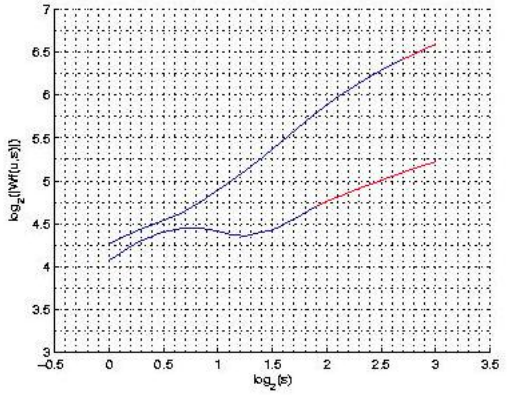
Table 2.2: Table with the estimated values for the decay of the wavelet-transform. $\Delta = \alpha + 1/2$ is the estimated decay with respect to $-\theta'(t)$ by using estimated scales for calculation. $\Delta' = \alpha + 1/2$ is the estimated decay with respect to $-\theta'(t)$ by using all scales $s \in [1, 6]$. $\Delta^{(2)} = \alpha + 1/2$ is the decay with respect to $\theta^{(2)}$.



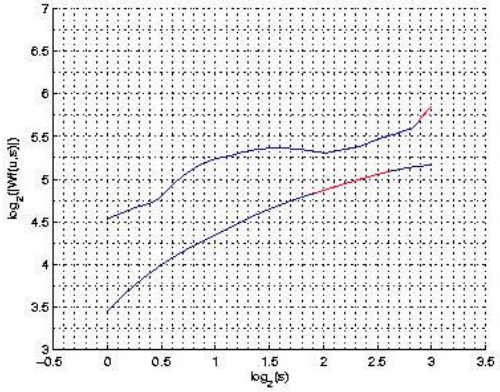
(a) Decay of the wavelet-transform at $\nu = 195$(upper) and $\nu = 43$(lower).



(b) Decay of the wavelet-transform at $\nu = 344$(upper) and $\nu = 469$(lower).



(c) Decay of the wavelet-transform at $\nu = 245$(upper) and $\nu = 376$(lower).



(d) Decay of the wavelet-transform at $\nu = 344$(upper) with $\theta^{(2)}$ and $\nu = 167$(lower) with θ'.

Figure 2.16: The slope of the wavelet-transform for some of the edges in the signal in Fig.(2.13b). The line segments marked with red, indicate the scales used to estimate the Lipschitz-regularity.

Chapter 3

2-D Wavelet-Transform.

The aim of the discussion in this chapter is to derive the 2-D wavelet-based trivial edge-detector. For the purpose of this thesis, such an edge-detector needs to fulfil two targets. Detect all the edges, and that the output can be further processed in order to locate the important edges in the signal. As will be discussed, these targets are fulfilled by using appropriately chosen 2-D wavelets.

In Sect.(3.1), 2-D wavelets and their wavelet-transform will be defined. 2-D wavelets will be defined as a collection of functions which satisfy the 2-D Admissibility-condition. One difference between the Admissibility-condition in 1-D and 2-D, is that the 2-D Admissibility-condition is allowed to vary with respect to (ω_x, ω_y). It will be discussed that all 2-D wavelets will detect the edges. If the Admissibility-condition is not constant, one can not directly use the output to locate the important edges. In this case the value of the wavelet-transform at an edge, it not invariant with respect to the orientation of the edge. For the purpose of this thesis, it would be sufficient to require the 2-D Admissibility-condition to be constant. The present definition has been chosen in order to illustrate one major obstacle for extending the theory into several dimensions.

The output of the wavelet-transform can not be used directly in order to detect the edges in a signal. In Sect.(3.2) it will be discussed how one can combine the (directional) wavelet-transforms defined in Sect.(3.1), to create an (undirectional) "wavelet-transform". This will be used in order to detect the edges in a 2-D function.

In the final section of this chapter the relation between the 2-D wavelet-transform and the 2-D Lipschitz-regularity will be investigated. This relation

will be used later in the thesis in order to construct edge-detectors which can be used to find the important edges in an image.

3.1 2-D Wavelet-Transform.

In this section 2-D wavelets and the 2-D wavelet-transform will be introduced.

The target of this thesis is to detect the objects of the US-image in Fig.(1.4) by using the 2-D wavelet-transform. To do this one has to find wavelets such that their respective wavelet-transform contains information which can be used to detect the edges in an image. One obstacle for introducing the wavelet-theory in several dimensions is that one has to take into account the orientation of the edges. Two problems emerge in 2-D because of orientation. First, one has to ensure that all the edges are be detected. This is necessary for function to be a 2-D wavelet. Second, one has to ensure that edges of similar strength and similar duration are represented with a similar value by the wavelet-transform. This is not necessarily true for all 2-D wavelets, but required if one wants to use the output to separate important from less important edges.

2-D wavelets will be defined such that there exist an inverse wavelet-transform. In Prop.(3.1.1) it will be proved that the definition of 2-D wavelets guarantees the existence of an inverse wavelet-transform. By the help of some examples, appropriate and less appropriate 2-D wavelets will be introduced.

The discussion begins with a definition of 2-D wavelets, and introducing some useful 2-D wavelets. Secondly, the 2-D wavelet-transform will be defined, and it will be verified that the definition of 2-D wavelets guarantee the existence of an inverse wavelet-transform.

Similarly as in the 1-D case, 2-D wavelets will be defined such that an inverse 2-D wavelet-transform exists. One difference is that a 2-D wavelet is not necessarily a single function, but may be composed of several functions.

Definition 3.1.1 (2-D Wavelets) *Assume* $\psi^k : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *and* $\psi^k \in L^2(\mathbb{R}^2)$ *for* $k = 1, ..., K$. *Then* $\psi(x, y) = \{\psi^1(x, y), ..., \psi^K(x, y)\}$ *is said to be a 2-D wavelet if there exists two constants* $0 < A \leq B < \infty$ *such that for all* $(\omega_x, \omega_y) \in \mathbb{R}^2 - \{(0, 0)\}$*;*

$$A \leq C_\psi(\omega_x, \omega_y) \leq B. \tag{3.1.1}$$

C_ψ is called the 2-D Admissibility-condition, and is given by;

$$C_\psi(\omega_x, \omega_y) = \int_0^\infty \left(\sum_{k=1}^{K} |\hat{\psi^k}(s\omega_x, s\omega_y)|^2 \right) \frac{ds}{s}.$$

The functions ψ^k will in this thesis be denoted semi-wavelets. Let $u = (u_1, u_2)$, $t = (t_1, t_2)$ and $s > 0$. Then

$$\psi_{u,s}(t) = \frac{1}{s}\psi\Big(\frac{t_1 - u_1}{s}, \frac{t_2 - u_2}{s}\Big),$$

denotes the scaled and translated wavelet. Let $\psi_{0,s} = \psi_s$.

□

Note: *The wavelet can also be defined with individual scales. Let $u = (u_1, u_2)$, $t = (t_1, t_2)$ and $s = (s_1, s_2)$.* Then

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s_1}\sqrt{s_2}}\psi\Big(\frac{t_1 - u_1}{s_1}, \frac{t_2 - u_2}{s_2}\Big)$$

denote the scaled and translated wavelet. This form of wavelets and the corresponding wavelet-transform will not be discussed in this thesis.

□

The next example will introduce some 2-D wavelets which will be important in this thesis.

Example 3.1.1 *Let*

$$\Theta(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} = \theta(x)\theta(y),$$

denote the 2-D Gaussian, where θ is the 1-D Gaussian. The functions corresponding to the partial derivative in the x-direction and the y-direction, denoted by resp. ψ^x and ψ^y, are up to sign given by;

$$\psi^x(x, y) = \frac{1}{2\pi} x e^{-\frac{x^2+y^2}{2}} = \psi^1(x)\theta(y)$$
$$\psi^y(x, y) = \frac{1}{2\pi} y e^{-\frac{x^2+y^2}{2}} = \theta(x)\psi^1(y).$$

ψ^1 is the 1-D wavelet corresponding to the 1.st derivative of the Gaussian. Then $\psi(x,y) = \{\psi^x(x,y), \psi^y(x,y)\}$ is a 2-D wavelet. By Prop.(B.0.3) and Prop.(B.0.1) it follows that;

$$\hat{\psi^x}(\omega_x,\omega_y) = \hat{\psi}(\omega_x)\hat{\theta}(\omega_y) = -(i\omega_x)\hat{\theta}(\omega_x)\hat{\theta}(\omega_y)$$
$$\hat{\psi^y}(\omega_x,\omega_y) = \hat{\theta}(\omega_x)\hat{\psi}(\omega_y) = -(i\omega_y)\hat{\theta}(\omega_x)\hat{\theta}(\omega_y).$$

For all $(\omega_x,\omega_y) \in \mathbb{R}^2 - \{(0,0)\}$;

$$\begin{aligned}
C_\psi(\omega_x,\omega_y) &= \int_0^\infty \left(|\hat{\psi^x}(s\omega_x,s\omega_y)|^2 + |\hat{\psi^y}(s\omega_x,s\omega_y)|^2\right)\frac{ds}{s} \\
&= \int_0^\infty \left(s^2\omega_x^2|\hat{\theta}(s\omega_x)\hat{\theta}(s\omega_y)|^2 + s^2\omega_y^2|\hat{\theta}(s\omega_x)\hat{\theta}(s\omega_y)|^2\right)\frac{ds}{s} \\
&= \int_0^\infty \left(s^2\omega_x^2|e^{-(s\omega_x)^2/2}e^{-(s\omega_y)^2/2}|^2 + s^2\omega_y^2|e^{-(s\omega_x)^2/2}e^{-(s\omega_y)^2/2}|^2\right)\frac{ds}{s} \\
&= \int_0^\infty \left(s^2\omega_x^2 e^{-(s\omega_x)^2}e^{-(s\omega_y)^2} + s^2\omega_y^2 e^{-(s\omega_x)^2}e^{-(s\omega_y)^2}\right)\frac{ds}{s} \\
&= \int_0^\infty \left(\omega_x^2+\omega_y^2\right)s e^{-s^2(\omega_x^2+\omega_y^2)}\,ds = \lim_{t\to\infty}\left[-\frac{1}{2}e^{-s^2(\omega_x^2+\omega_y^2)}\right]_{s=0}^{s=t} \\
&= \frac{1}{2} - \lim_{s\to\infty}\frac{1}{2}e^{-s^2(\omega_x^2+\omega_y^2)} = 1/2.
\end{aligned}$$

By Def.(3.1.1), $\psi(x,y) = \left\{\psi^x(x,y), \psi^y(x,y)\right\}$ is a 2-D wavelet.

Assume $\mathbf{n} = (n_1,n_2)$ and $\mathbf{m} = (m_1,m_2)$ are unit-vectors in $\mathbb{R}^2$, and let $\psi^{\mathbf{n}} = \frac{-\partial\Theta}{\partial\mathbf{n}}$ and $\psi^{\mathbf{m}} = \frac{-\partial\Theta}{\partial\mathbf{m}}$. Let $\psi'(x,y) = \left\{\psi^{\mathbf{n}}(x,y), \psi^{\mathbf{m}}(x,y)\right\}$. Note that;

$$\psi^{\mathbf{n}} = \frac{\partial\Theta}{\partial\mathbf{n}} = \mathbf{n}\cdot\left(\psi^x,\psi^y\right) \;\Rightarrow\; \hat{\psi}^{\mathbf{n}} = \mathbf{n}\cdot\left(\hat{\psi}^x,\hat{\psi}^y\right).$$

Assume that the direction of $\mathbf{n}$ and $\mathbf{m}$ to the x-axis is given by α and β, and that the direction of the vector (ω_x,ω_y) to the x-axis is γ. Note that α and β are fixed. Note that $\mathbf{A}\cdot\mathbf{B} = |\mathbf{A}||\mathbf{B}|\cos(\phi)$ where $\cdot$ is the dot-product, and ϕ in the angle between the vectors $\mathbf{A}$ and $\mathbf{B}$. By the computations in the previous

example, it follows that;

$$\begin{aligned} C_{\psi'}(\omega_x,\omega_y) &= \int_{-\infty}^{\infty} \left(|\hat{\psi}^x(s\omega_x,s\omega_y)|^2 + |\hat{\psi}^y(s\omega_x,s\omega_y)|^2\right) \frac{ds}{s} \\ &= \int_{-\infty}^{\infty} \left(|\hat{\theta}(s\omega_x,s\omega_y)|^2 s^2(|\mathbf{n}\cdot(\omega_x,\omega_y)|^2 + |\mathbf{m}\cdot(\omega_x,\omega_y)|^2)\right) \frac{ds}{s} \\ &= \int_{-\infty}^{\infty} |\hat{\theta}(s\omega_x,s\omega_y)|^2 s(\omega_x^2+\omega_y^2)\, ds \left(\cos^2(\alpha-\gamma)+\cos^2(\beta-\gamma)\right) \\ &= \frac{1}{2}\Big(\cos^2(\alpha-\gamma)+\cos^2(\beta-\gamma)\Big) \end{aligned}$$

First one may note that if $\mathbf{n} \parallel \mathbf{m} \Rightarrow \alpha = \beta$*, then there exists infinitely many* $(\omega_x,\omega_y) \in \mathbb{R}^2 - \{(0,0)\}$ *such that* $C_{\psi'}$ *equals* 0*. The second observation is that if* $\mathbf{n} \perp \mathbf{m} \Rightarrow \beta = \alpha + \pi/2$*, then* $C_{\psi''} = 1/2$*. Finally, one may observe that if there exists* $\delta > 0$ *such that* $\delta < \alpha - \beta$*, then the 2-D Admissibility-condition is satisfied. Given two vectors* $\mathbf{n} \nparallel \mathbf{m}$*, then Def.(3.1.1) gives that* $\psi' = \{\psi^{\mathbf{n}}, \psi^{\mathbf{m}}\}$ *is a 2-D wavelet.*

Let ψ^2 *denote the 1-D Mexican-Hat. With a similar approach as above, one may verify that* $\psi''(x,y) = \{\psi^2(x)\theta(y), \theta(x)\psi^2(y), 2\psi^1(x)\psi^1(y)\}$ *and* $\psi'''(x,y) = \{\psi^2(x)\theta(y), \theta(x)\psi^2(y)\}$ *are 2-D wavelets, with resp.* $C_{\psi''} = 1/2$ *and* $1/4 \leq C_{\psi'''}(\omega_x,\omega_y) = 1/2 - \frac{\omega_x^2\omega_y^2}{(\omega_x^2+\omega_y^2)^2} \leq 1/2$.

□

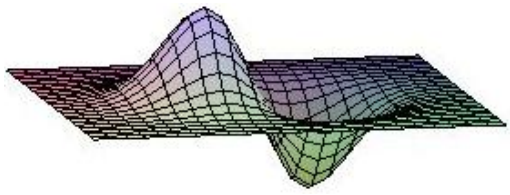


Figure 3.1: A 2-D semi-wavelet corresponding to the partial derivative of the 2-D Gaussian.

In Ex.(3.1.2) it will be explained from an geometric viewpoint, what goes wrong if $C_\psi(\omega_x,\omega_y) = 0$ for $(\omega_x,\omega_y) \neq (0,0)$. The first two 2-D wavelet in Ex.(3.1.1) are suitable for detecting step-edges, and the last two wavelets

can be used to detect angle-edges. However, the second and the last wavelet are not suitable for detecting edges. It will detect all edges, but it is not invariant of the direction of the edge. This means that edges of equal strength and duration may be given different values of the wavelet-transform. These statements will be discussed in the next section.

Similarly as with 2-D wavelets, the 2-D Continuous Wavelet-Transform comes in pairs.

Definition 3.1.2 (2-D Wavelet-transform) *Assume* $f \in L^2(\mathbb{R}^2)$ *and* $\psi = \{\psi^1, ..., \psi^K\}$ *is a 2-D wavelet. Define for each of the semi-wavelets, the 2-D Continuous (semi-)Wavelet-Transform as;*

$$W^k f((u,v),s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{1}{s}\psi^k\Big(\frac{x-u}{s}, \frac{y-v}{s}\Big)\, dxdy, \tag{3.1.2}$$

for $(u,v) \in \mathbb{R}^2$ *and* $s > 0$. *The 2-D Continuous Wavelet-Transform is defined as the collection* $Wf((u,v),s) = \{W^1 f((u,v),s), ..., W^K f((u,v),s)\}$, *for* $(u,v) \in \mathbb{R}^2$ *and* $s > 0$.

□

The 2-D wavelet-transform can be expressed as a collection of 2-D convolution-products,

$$Wf((u,v),s) = \Big\{f * \tilde{\psi}^1(u,v), ..., f * \tilde{\psi}^K(u,v)\Big\}, \tag{3.1.3}$$

where $\tilde{\psi}(x,y) = \psi(-x,-y)$. For the 2-D wavelet $\psi = \{\psi^x, \psi^y\}$, the wavelet-transform can be expressed as;

$$Wf((u,v),s) = \Big\{s\frac{\partial}{\partial x}(f * \Theta_s(u,v)), s\frac{\partial}{\partial y}(f * \Theta_s(u,v))\Big\} \tag{3.1.4}$$

The next question is whether one can recover a 2-D signal by its 2-D wavelet-transform. Reconstruction of signals will not be discussed in this thesis, but is a typical extension of edge-detection. The machinery of edge-detection and reconstruction will allow one to manipulate images, e.g. to remove noise in images. For simplicity the statement is proved with respect to a 2-D wavelet which is constituted of a pair of semi-wavelets.

Proposisjon 3.1.1 *Assume* $\psi = \{\psi^a, \psi^b\}$ *is 2-D wavelet such that;*

$$A < C_\psi(\omega_x, \omega_y) = \int_0^\infty \left(|\hat{\psi^a}(s\omega_x, s\omega_y)|^2 + |\hat{\psi^b}(s\omega_x, s\omega_y)|^2 \right) \frac{ds}{s} < B,$$

for some constants $0 < A < B < \infty$. *Any* $f \in L^2(\mathbb{R}^2) \cap L^1(\mathbb{R}^2)$ *satisfies;*

$$f(x,y) = \frac{1}{C_\psi(\omega_x, \omega_y)} \int_0^\infty \left(\iint_{-\infty}^\infty W^a f((u,v),s) \frac{1}{s} \psi^a\Big(\frac{x-u}{s}, \frac{y-v}{s}\Big)\, dxdy \right. \tag{3.1.5}$$

$$\left. + \iint_{-\infty}^\infty W^b f((u,v),s) \frac{1}{s} \psi^b\Big(\frac{x-u}{s}, \frac{y-v}{s}\Big)\, dxdy \right) \frac{ds}{s^3}.$$

Proof: The idea of the proof is to show that the right-side, denoted $b(x,y)$, of Eq.(3.1.5) equals $f(x,y)$. First one may observe that the two inner integrals equal a convolution-product. I.e.;

$$b(x,y) = \frac{1}{C_\psi(\omega_x, \omega_y)} \int_0^\infty \left(W^a f((\cdot,\cdot),s) * \psi^a_s(x,y) + W^b f((\cdot,\cdot),s) * \psi^b_s(x,y) \right) \frac{ds}{s^3}.$$

By Eq.(3.1.3) it follows that;

$$b(x,y) = \frac{1}{C_\psi(\omega_x, \omega_y)} \int_0^\infty \left(f * \tilde{\psi^a_s} * \psi^a_s(x,y) + f * \tilde{\psi^b_s} * \psi^b_s(x,y) \right) \frac{ds}{s^3}.$$

By Prop.(B.0.6) and Prop.(B.0.4), the 2-D Fourier-Transform of $b(x,y)$, can be expressed as;

$$\begin{aligned}
\hat{b}(\omega_x, \omega_y) &= \frac{1}{C_\psi(\omega_x, \omega_y)} \int_0^\infty \left(\hat{f}(\omega_x, \omega_y) s \bar{\hat{\psi}}^a(\omega_x, \omega_y) s \hat{\psi^a}(s\omega_x, s\omega_y) \right. \\
&\quad \left. + \hat{f}(\omega_x, \omega_y) s \bar{\hat{\psi}}^b(s\omega_x, s\omega_y) s \hat{\psi^b}(s\omega_x, s\omega_y) \right) \frac{ds}{s^3} \\
&= \frac{\hat{f}(\omega_x, \omega_y)}{C_\psi(\omega_x, \omega_y)} \int_0^\infty \left(|\hat{\psi^a}(\omega_x, \omega_y)|^2 + |\hat{\psi^b}(s\omega_x, s\omega_y)|^2 \right) \frac{ds}{s} \\
&= \hat{f}(\omega_x, \omega_y).
\end{aligned}$$

Since $\hat{b}(\omega_x, \omega_y) = \hat{f}(\omega_x, \omega_y)$, it follows that $b(x,y) = f(x,y)$.

□

The next example shows a calculation of the 2-D wavelet-transform, and illustrates the importance of the direction of the semi-wavelet used by the wavelet-transform.

Example 3.1.2 *Assume $f(x,y) = u(x)$ where u is the Heaviside-function, and $\psi^x(x,y) = \psi(x)\theta(y)$ and $\psi^y(x,y) = \theta(x)\psi(y)$. Then for $(w,v) \in \mathbb{R}^2$;*

$$\begin{aligned} W^x f((w,v),s) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{1}{s}\psi(\frac{x-w}{s})\theta(\frac{y-v}{s})\,dxdy \\ &= \int_{-\infty}^{\infty}\theta(\frac{y-v}{s})\,dy\int_{-\infty}^{\infty}u(x)\frac{1}{s}\psi(\frac{x-w}{s})\,dx \\ &= \sqrt{s}\int_{-\infty}^{\infty}\theta(t)\,dt\int_{0}^{\infty}\frac{1}{\sqrt{s}}\psi(\frac{x-w}{s})\,dx \\ &= \sqrt{s}\int_{0}^{\infty}\frac{1}{\sqrt{s}}\psi(\frac{x-w}{s})\,dx = s^{3/2}\theta_s(w), \end{aligned}$$

with change of variable $y - v = ts$, and the final equality follows from Ex.(2.1.2). On the other hand;

$$\begin{aligned} W^y f((w,v),s) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{1}{s}\theta(\frac{x-w}{s})\psi\frac{y-v}{s})\,dxdy \\ &= \int_{-\infty}^{\infty}u(x)\theta(\frac{x-w}{s})\,dy\int_{-\infty}^{\infty}\frac{1}{s}\psi(\frac{y-v}{s})\,dx = 0, \end{aligned}$$

This follows since all 1-D wavelets have 0-mean. This latter computation illustrate why ψ^y (or ψ^x) alone can not be a 2-D continuous wavelet. There are an infinitely number of functions $g(x,y)$ such that $Wg((w,v),s) = 0$, $\forall(w,v) \in \mathbb{R}^2$, e.g. all functions of the form $g(x,y) = Au(x-\tau)$ for $A, \tau \in \mathbb{R}$. This implies that there can not exist a unique inverse.

For the function given by $h(x,y) = \delta(x)\delta(y)$ the semi-wavelet-transform with respect to $\psi^x(x,y)$ is given by,

$$W^x h((w,v),s) = \Big(\delta(x) * \tilde{\psi}_s(x-w)\Big)\Big(\delta(y) * \tilde{\theta}_s(x-v)\Big) = \tilde{\psi}_s(x)\tilde{\theta}_s(y) = \tilde{\psi}^x(x,y).$$

$\tilde{f}(t) = f(-t)$ *and* $\tilde{g}(x,y) = g(-x,-y)$.

Similarly one may verify that for the function $g(x,y) = \delta(x)$, then;

$$W^x g((w,v),s) = \tilde{\psi}^1(x)$$
$$W^y g((w,v),s) = 0.$$

□

The question is how one can detect edges from the 2-D wavelet-transform. As discussed in the latter example, ψ^x does not detects horizontal step-edges and ψ^y does not detect vertical step-edges. On the other hand, they both detect diagonal step-edges. The next section will discuss how one can combine the two semi-wavelets to easily detect all the edges in a signal.

3.2 Wavelet-Transform and Edges in $\mathbb{R}^2$.

In this section it will be discussed how one can detect edges in an image by using the 2-D CWT. At the end of the section the wavelet-based trivial edge-detector for images will be presented.

As discussed in Sect.(1.2.1), the first step of detecting the "important" edges in a signal is first to find all the edges in the signal. Detecting edges in a 2-D signal is not as straight forward as in 1-D case. As indicated by Ex.(3.1.2), an edge may not be "felt" by the wavelet-transform. It the edge is not felt by the wavelet-transform, it can not be detected. If the output of the 2-D wavelet-based trivial edge-detector should be of any use, one has to be sure that all edges are detected, and equally important, that similar edges are given a similar value by the wavelet-transform.

The target of this section is to derive a "wavelet-transform" which is invariant of the direction of the edge. Such a "wavelet-transform" should fulfil three criteria. First, all edges should be detected. This is fulfilled if the wavelet satisfies the 2-D admissibility-condition. Second, similar edges should be represented with an equal value by the transform. This criteria is fulfilled if the admissibility-condition is constant for all $(\omega_x, \omega_y) \in \mathbb{R}^2 - \{(0,0)\}$. In addition one should be able to reconstruct the signal. A natural starting point for such an investigation is the wavelet-transform defined in the previous section. The question is how one can use the (semi-)wavelet-transforms to find a representation of the signal which satisfies the three criteria.

The investigation begins with considering how one can find points of maximal-change in smooth functions. From this discussion, it will be described how one can combine ψ^x and ψ^y and their respective wavelet-transforms to find the step-edges in an arbitrary function. At the end, it will be discussed how one can detect angle-edges in a signal. From this discussion it will be derived that the choice of 2-D wavelet should not only depend on whether all edges are detected. I.e. it is desirable to use a 2-D wavelet for which the admissibility-condition is constant.

One question which needs to be answered is how one would like to detect edges in 2-D. Similarly as in 1-D, one can detect edges both by tracing local maxima, and by tracing zero-crossings. For instance the Marr and Hildreth edge-detector in [8] uses zero-crossings to detect the edges. It is easier to obtain information about the strength of an edge by using the maxima-approach. Therefore the maxima-approach has been used in this thesis.

Assume $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ is continuously differentiable in $\mathbb{R}\times\mathbb{R}$. In college-math one learn that at a point (x_0, y_0), the direction of maximal-change is given by the vector $\big(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0)\big)$. The direction in which f changes maximally can be expressed as;

$$\begin{cases} \tan^{-1}\left(\frac{\frac{\partial f}{\partial y}(x_0,y_0)}{\frac{\partial f}{\partial x}(x_0,y_0)}\right) & \text{if } \frac{\partial f}{\partial x}(x_0, y_0) \geq 0 \\ \pi - \tan^{-1}\left(\frac{\frac{\partial f}{\partial y}(x_0,y_0)}{\frac{\partial f}{\partial x}(x_0,y_0)}\right) & \text{if } \frac{\partial f}{\partial x}(x_0, y_0) < 0\ . \end{cases}$$

The length of the vector at (x_0, y_0), denoted $M(x_0, y_0)$, is given by;

$$M(x_0, y_0) = \sqrt{\Big(\frac{\partial f}{\partial x}(x_0, y_0)\Big)^2 + \Big(\frac{\partial f}{\partial y}(x_0, y_0)\Big)^2}.$$

The function $f(x, y)$ has locally maximal growth at (x_0, y_0) if $M(x, y)$ is locally maximal at (x_0, y_0) in the direction of maximal change. As in 1-D, the idea is to imitate this procedure to general functions f.

Prop.(B.0.7) verify that $(f * \Theta)$ is continuously differentiable. One can therefore use the procedure above to find the points where $(f * \Theta)$ has maximal-growth . One is not guaranteed these points coincide with the position of the edges in f, or if this procedure will detect all edges, or if artificial edges are created. In this thesis it will be assumed that Θ is a suitable function for the purpose of this thesis.

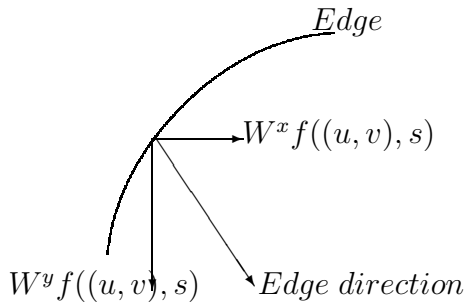


Figure 3.2: Illustration of how one locally estimate the direction of an edge in $\mathbb{R}^2$.

As verified by Eq.(3.1.4), the CWT with respect to the semi-wavelets ψ^x and ψ^y can be considered as the partial derivative of $s(f * \Theta_s(u,v))$ in resp. the x-direction and y-direction. $W^x f((u,v),s)$ correspond to how much $f * \Theta$ changes in the horizontal-direction and $W^y f((u,v),s)$ to how much $f * \Theta$ changes in the vertical-direction. The next question is how one can combine these to detect all the edges in a signal. The idea is that instead of using a semi-wavelet with a fixed orientation for the entire signal, one can at each point compute the semi-wavelet-transform in the direction which the signal changes the most. I.e. if the signal at a point changes most in the direction given by $\mathbf{n}$, then the semi-wavelet-transform at that point should be computed with respect to the semi-wavelet $\psi^{\mathbf{n}}$. The direction of maximal change of $(f * \Theta)$ at a point (u,v) is given by;

$$Af((u,v),s) = \begin{cases} \alpha(u,v) & \text{if } W^1 f((u,v),s) \geq 0 \\ \pi - \alpha(u,v) & \text{if } W^1 f((u,v),s) < 0 \end{cases},$$

where

$$\alpha(u,v) = \tan^{-1}\left(\frac{W^2 f((u,v),s)}{W^1 f((u,v),s)}\right).$$

This give the direction $\mathbf{n}(u,v) = (\cos Af(u,v), \sin Af(u,v))$ for which the 2-D CWT should be computed at (u,v). The absolute value of $W^{\mathbf{n}} f((u,v),s)$ is given by;

$$|W^{\mathbf{n}} f((u,v),s)| = |s\frac{\partial}{\partial \mathbf{n}}\Big(f * \Theta\Big)| = |s\nabla(f * \Theta) \cdot \mathbf{n}| = Mf((u,v),s)$$

where;

$$Mf((u,v),s) = \sqrt{W^x f((u,v),s)^2 + W^y f((u,v),s)^2}.$$

$Mf((u,v),s)$ is called the modulus of f. It follows that it is sufficient to compute the 2-D wavelet-transform with respect to the 2-D wavelet $\psi = \{\psi^x, \psi^y\}$. The candidate edges are found among the points where $Mf((u,v),s)$ is locally maximum in the direction given by $\mathbf{n}(u,v) = (\cos Af((u,v),s), \sin Af((u,v),s))$. The relations;

$$W^x f((u,v),s) = Mf((u,v),s)\cos Af((u,v),s) \tag{3.2.1}$$
$$W^y f((u,v),s) = Mf((u,v),s)\sin Af((u,v),s), \tag{3.2.2}$$

verify that one can reconstruct the signal from Mf.

The following summarizes the *trivial edge-detector* in 2-D.

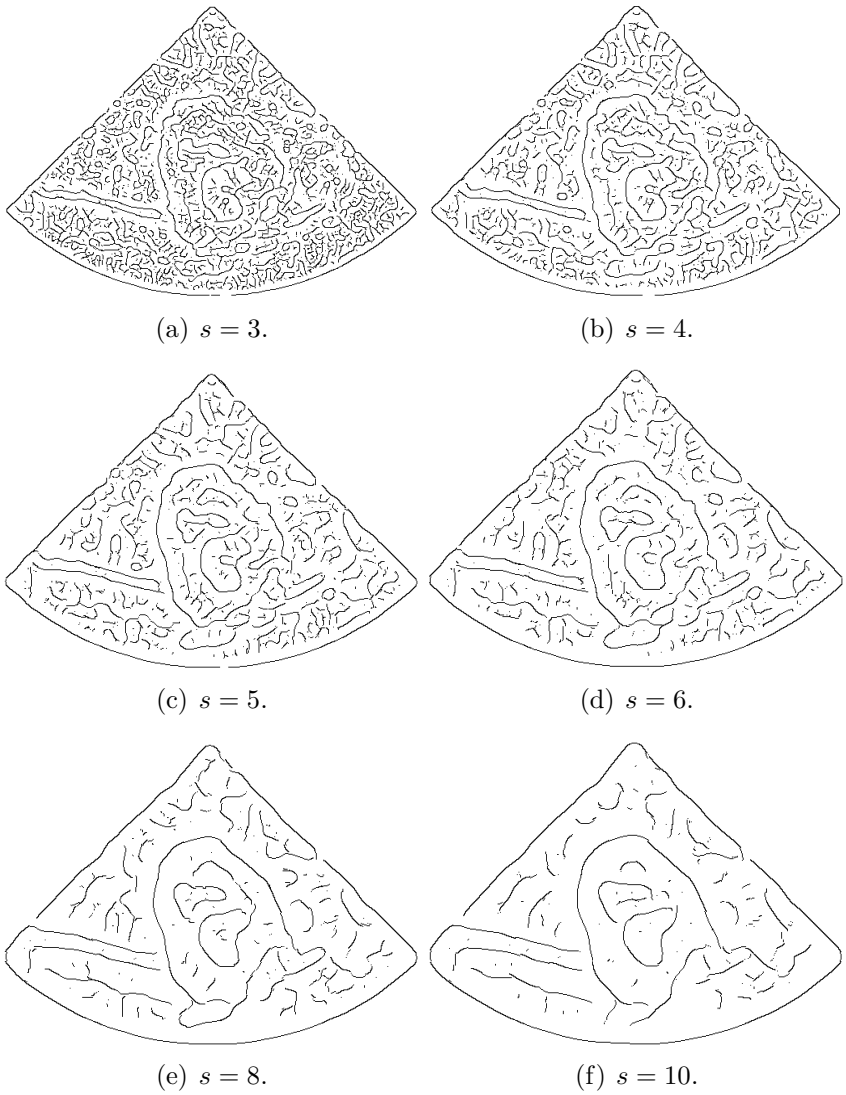
1. Compute $W^x f((u,v),s)$ and $W^y f((u,v),s)$ for $s \in \{s_1, ..., s_N\}$.
2. Compute $Mf((u,v),s)$.
3. Compute $Af((u,v),s)$.
4. Find the modulus-maximum of $Mf((u,v),s)$ in the local direction given by $\mathbf{n}(u,v)$.

The output of the trivial edge-detector is the set of all modulus-maximum at all the scales where the CWT is computed;

$$\{((x_i^m, y_i^m), s_m)\} \quad \text{for } s_m \in \{s_1, ..., s_N\}.$$

To round off the section it will be briefly explained how angle-edges can be detected in $\mathbb{R}^2$. This will demonstrate the difference between the 2-D wavelets ψ'' and ψ''' in Ex.(3.1.1). Assume $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ has an angle-edge at $(0,0)$ in the direction given by $\mathbf{n} = (n_1, n_2)$. This imply that $g = \frac{\partial f}{\partial \mathbf{n}}$ has a step-edge at $(0,0)$ in the direction given by $\mathbf{n}$. From the previous discussion, step-edges can be detected by the CWT with respect to the wavelet $\psi^{\mathbf{n}}$. I.e. for $u = (u_1, u_2)$ and $s > 0$;

$$\begin{aligned}|W^{\mathbf{n}} g(u,s)| &= |s\nabla(g * \Theta_s)| \\ &= \left(\left(W^a f(u,s) + W^c f(u,s)\right)^2 + \left(W^b f(u,s) + W^c f(u,s)\right)^2 \right)^{1/2}.\end{aligned}$$



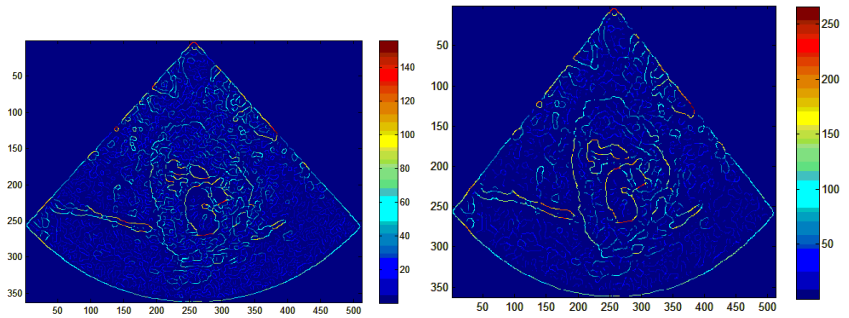
(a) $s = 3$. (b) $s = 4$.

(c) $s = 5$. (d) $s = 6$.

(e) $s = 8$. (f) $s = 10$.

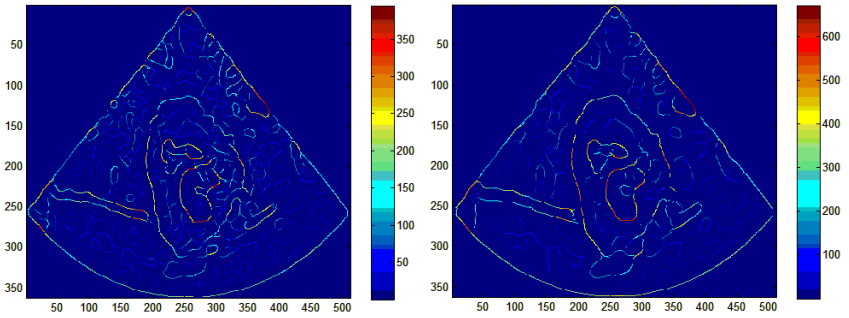
Figure 3.3: The output of the 2-D wavelet-based trivial edge-detector applied to the signal in Fig.(1.4). The output does only take into account the position of the modulus-maximum.

$W^a f$, $W^b f$ and W^c are the semi-wavelet-transform computed with the semi-wavelets $\psi^a(x, y) = \psi^2(x)\theta(y)$, $\psi^b(x, y) = \theta(x)\psi^2(y)$ and $\psi^c(x, y) = \psi^1(x)\psi^1(y)$. From this one may observe the difference between the wavelets ψ'' and ψ'''.



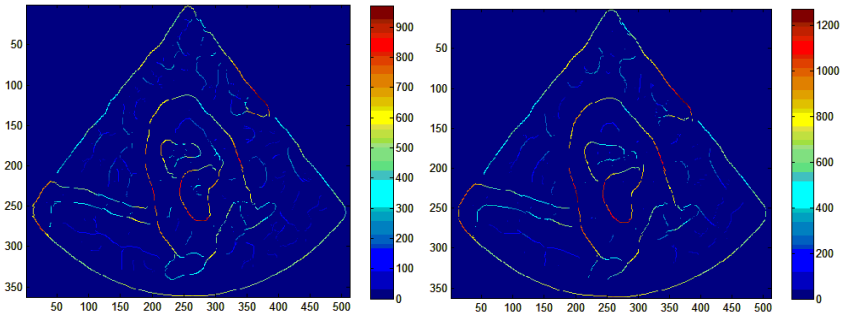
(a) $s = 2$.

(b) $s = 3$.



(c) $s = 4$.

(d) $s = 6$.



(e) $s = 8$.

(f) $s = 10$.

Figure 3.4: The output of the 2-D wavelet-based trivial edge-detector applied to the signal in Fig.(1.4). The output illustrates the strength of the modulus-maximum.

The modulus of the semi-wavelets of ψ''', will be less for a step-edge which is not horizontal or vertical. With respect the wavelet ψ'', the value of the

wavelet-transform does not depend on the direction of the edge. One should note that the detection-performance does not differ between the wavelets, i.e. they detect the same edges. It is a necessity that edges of similar strength are given a similar value by the wavelet-transform. Otherwise one could not use the amplitude of the modulus-maximum in order to separate the important edges in the signal.

3.3 2-D Wavelet-Transform and Lipschitz-Regularity in $\mathbb{R}^2$.

In this section the Lipschitz-regularity of 2-D functions, and its relation with the 2-D continuous wavelet-transform will be established.

The relation between the wavelet-transform and the Lipschitz-regularity will allow one to distinguish between different kinds of edges, such as Dirac-edges, step-edges and angle-edges. This will enable to separate different features in an image.

The discussion of Lipschitz-regularity and wavelet-transform will be restricted to step-edges, i.e. for $\alpha \leq 1$. First, the definition of 2-D Lipschitz-regularity will be presented, and a proof which relates the 2-D Lipschitz-regularity at a point to the 1-D Lipschitz-regularity of 1-D rays which go through the point. Global and local conditions for estimating the Lipschitz-regularity by the wavelet-transform across scales will be discussed.

The obstacle in 2-D is that a point may be smooth in one direction and be singular in another direction. The Lipschitz-regularity is defined to estimate the worst behaviour in the proximity of the point, i.e. the Lipschitz-regularity measure the smoothness in the direction of worst behaviour.

Definition 3.3.1 (2-D Lipschitz Conditions, Regularity and Singularity [1])

Local Lipschitz *Let $0 \leq \alpha \leq 1$. A function $f(x, y)$ is said to be pointwise Lipschitz-α at (x_0, y_0) if and only if there exist an $A \in \mathbb{R}_+$, an $\epsilon_k, \epsilon_h > 0$ such that for all $|k| < \epsilon_k$ and $|h| < \epsilon_h$;*

$$|f(t_0 + k, t_0 + h) - f(t_1, t_1)| \leq A(h^2 + k^2)^{\alpha/2}.$$

Global Lipschitz *Let* $]a,b[\times]c,d[$ *be an open interval and* $0 \leq \alpha \leq 1$. *Then* $f(x,y)$ *is said to be globally Lipschitz-*α *over* $]a,b[\times]c,d[$ *if and only if there exist an* $A \in \mathbb{R}$ *and* $\epsilon_k, \epsilon_h > 0$ *such that for any* $(t_0,t_1) \in]a,b[\times]c,d[$ *and* $t_0+k \in]a,b[, t_1+h \in]c,d[$*;*

$$|f(t_0+k,t_1+h) - f(t_0,t_1)| \leq A(h^2+k^2)^{\alpha/2}.$$

Lipschitz Regularity *The Lipschitz-regularity of* $f(x,y)$ *locally at* t_0 *or globally over* $]a,b[\times]c,d[$ *is the superior bound of all* α*-values such that* $f(x,y)$ *is Lipschitz-*α *at* (t_0,t_1) *or over* $]a,b[\times]c,d[$.

The following statement may be useful for establishing the local Lipschitz-regularity at a point.

Statement 3.3.1 *Assume* $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$. *Define the function* $g_\omega : [-\epsilon,\epsilon] \to \mathbb{R}$ *by;*

$$g_\omega(t) = f(t_0 + t\cos(\omega), t_1 + t\sin(\omega)),$$

for $\omega \in [0,\pi[$ *and* $\epsilon > 0$. *The local Lipschitz-regularity of* f *at* (t_0,t_1) *is less or equal the local Lipschitz-regularity of* g_ω *at* 0 *for all* $\omega \in [0,\pi[$.

Proof: Assume that the local Lipschitz-regularity of $f(x,y)$ equals α at (t_0,t_1). By Def.(3.3.1) there exists an $A > 0, \epsilon_h, \epsilon_k > 0$ such that for all $|h| < \epsilon_h$ and $|k| < \epsilon_k$;

$$|f(t_0+k,t_1+h) - f(t_0,t_1)| \leq A(h^2+k^2)^{\alpha/2}.$$

This inequality holds for $|h| < \epsilon_h$ and $|k| < \epsilon_k$, and therefore holds for all $h = \delta\cos\omega$ and $k = \delta\sin\omega$ for all $\omega \in [0,\pi[$ with $|\delta| < \min\{\epsilon_h,\epsilon_k\}$. Therefore;

$$\begin{aligned} |f(t_0+\delta\cos\omega, t_1+\delta\sin\omega) - f(t_0,t_1)| &\leq A(\delta^2)^{\alpha/2} \\ &\Downarrow \\ |g_\omega(\delta) - g_\omega(0)| &\leq A(\delta^2)^{\alpha/2} = A|\delta|^\alpha. \end{aligned}$$

Def.(2.2.1) gives that g_ω is locally Lipschitz-α at 0 for all ω.

□

The next example will use this statement to prove that the local Lipschitz-regularity of $f(x,y) = u(x)$, where $u(x)$ is the Heaviside, equals 0 for any $(t_0,t_1) = (0,t_1), t_1 \in \mathbb{R}$.

Example 3.3.1 *Assume $f(x,y) = u(x) \cdot 1$, and that $(x_0, y_0) = (0,0)$. All 1-D rays which go through (x_0, y_0) equals the Heaviside if the ray is not perpendicular to the x-axis, and equal 1 if the ray is perpendicular to the x-axis. Ex.(2.2.1) and Stat.(3.3.1) give that the Lipschitz-regularity of f is less or equal 0 at 0. On the other hand, by choosing $A = 2$, $\epsilon_h, \epsilon_k = 1/2$ it follows that for all $|h| < \epsilon_h, |k| < \epsilon_k$;*

$$|f(k,h) - f(0,0)| = \begin{cases} 0 & \text{if } k \geq 0 \\ 1 & \text{if } k < 0 \end{cases} < 2(h^2 + k^2)^{0/2} = 2.$$

A similar computation holds for any $t_1 \in \mathbb{R}$.

□

In the previous chapter, a definition was used in order to extend the concept of Lipschitz-regularity for negative values for α. A similar definition will not be investigated in 2-D. It will be assumed that Dirac-edges are Lipschitz-(-1).

As in 1-D,there exists a relation between the decay of the wavelet-transform and the Lipschitz-regularity of a function. This decay is controlled by $Mf(u,s)$. The next theorem relate $Mf(u,s)$ with the uniform Lipschitz-regularity.

Theorem 3.3.1 (Global Lipschitz-regularity.[4]) *Assume $0 < \alpha < 1$. The function $f \in L^2(\mathbb{R}^2)$ is globally Lipschitz α in $]a,b[\times]c,d[$ if there exist a constant $A \in \mathbb{R}_+$ such that for all $(t_0, t_1) \in]a,b[\times]c,d[$,*

$$|Mf((t_0,t_1),s)| \leq As^{\alpha+1}.$$

□

This theorem proves that the global Lipschitz-regularity can be measured by the decay of the $Mf(u,s)$ across scales, by using an equation similar to Eq.(2.4.5).

Proposisjon 3.3.1 (Local Lipschitz-regularity) *Assume $f \in L^2(\mathbb{R}^2)$ and $\psi(x,y) = \{\psi^x(x,y), \psi^y(x,y)\}$. If there exists an $0 < \alpha < 1$, a constant A and an $\alpha' < \alpha$ such that;*

$$\forall((u,v),s) \in \mathbb{R}^2 \times \mathbb{R}\ ,\ |Mf(u,s)| \leq As^{\alpha+1}\left(1 + \left(\left|\frac{u-t_0}{s}\right| + \left|\frac{v-t_1}{s}\right|\right)^{\alpha'}\right),$$

then f is Lipschitz-α at (t_0, t_1).

Before this proposition is proved, note that there exists an A, $\epsilon_k, \epsilon_h > 0$ such that equation in Def.(3.3.1) is satisfied, if and only is there exists $\epsilon'_k, \epsilon'_h > 0$ such that for all $h' < \epsilon'_h, k' < \epsilon'_k$;

$$|f(t_0 + h', t_1 + k') - f(t_0, t_1)| \leq \big(|h'| + |k'|\big)^{\alpha}.$$

In addition note that the semi-wavelets ψ^x and ψ^y and their partial derivatives, all satisfy that for any $m \in \mathbb{N}$ there exists a constant C_m such that;

$$\forall\ (x, y) \in \mathbb{R}^2,\ |\psi(x, y)| \leq \frac{C_m}{1 + (|x| + |y|)^m}. \tag{3.3.1}$$

The idea of the proof is simple; to verify that f satisfied Def.(3.3.1). Instead of using f which one do not know anything about, one can use the reconstruction-formula as a starting-point. With respect to the variable s one can divide the integral in Prop.(3.1.1) into two integrals; $\int_0^\infty = \int_0^a + \int_a^\infty$. By Properties of convolution, the integral $\int_a^\infty$ is smooth, i.e. discontinuities in the reconstruction of f will first appear as $a \to 0$. By establishing decay conditions for small a, one can derive the desired result.

Proof of Prop.(3.3.1): The 2-D wavelet reconstruction formula proves that $f(x, y)$ can be written on the form;

$$f(x, y) = \sum_{j=-\infty}^{\infty} \Delta_j(x, y),$$

with

$$\Delta_j(x, y) = \frac{1}{C_\psi} \int_{2^j}^{2^{j+1}} \left(\iint_{-\infty}^{\infty} W^x f((u, v), s) \frac{1}{s} \psi^x\Big(\frac{x - u}{s}, \frac{y - v}{s}\Big)\, dxdy \right.$$
$$\left. + \iint_{-\infty}^{\infty} W^y f((u, v), s) \frac{1}{s} \psi^y\Big(\frac{x - u}{s}, \frac{y - v}{s}\Big)\, dxdy \right) \frac{ds}{s^3}.$$

To prove that $f(x, y)$ is Lipschitz-α at (t_0, t_1), one needs an constant C such that;

$$|f(t_0 + h, t_1 + k) - C| \leq A(|h| + |k|)^{\alpha}.$$

The idea is to prove that the constant $\sum_{j=-\infty}^{\infty} \Delta_j(t_0, t_1)$ satisfies this condition. One necessary condition is that this sum is finite. The first step is to

prove that $|\Delta_j(x,y)|$ is bounded for all $j \in \mathbb{Z}$ and decays as $|j| \to \infty$. Recall Eq.(3.2.1), Eq.(3.2.2) and the asymptotic decay condition of Eq.(3.3.1). It follows that;

$$\begin{aligned}
\Delta_j(x,y) &= \frac{1}{C_\psi}\left| \int_{2^j}^{2^{j+1}} \left(\iint_{-\infty}^{\infty} Mf((u,v),s)\Big\{ \frac{\cos Af((u,v),s)}{s}\psi^x\Big(\frac{x-u}{s},\frac{y-v}{s}\Big) \right.\right. \\
&\quad \left.\left. + \frac{\sin Af((u,v),s)}{s}\psi^y\Big(\frac{x-u}{s},\frac{y-v}{s}\Big)\Big\}\, dudv \right)\frac{ds}{s^3}\right| \\
&\le \frac{1}{C_\psi}\int_{2^j}^{2^{j+1}} \left(\iint_{-\infty}^{\infty} As^\alpha\Big(1+\Big(\Big|\frac{u-t_0}{s}\Big|+\Big|\frac{v-t_1}{s}\Big|\Big)^{\alpha'}\Big) \right. \\
&\quad \left. \Big\{ \frac{C_m}{1+\big(|(x-u)/s|+|(y-v)/s|\big)^m} + \frac{C'_m}{1+\big(|(x-u)/s|+|(y-v)/s|\big)^m}\Big\}\, dudv \right)\frac{ds}{s^2} \\
&\le \frac{A\cdot\max\{2C'_m,2C_m\}}{C_\psi}\int_{2^j}^{2^{j+1}} \left(\iint_{-\infty}^{\infty} s^\alpha\Big(1+\Big(\Big|\frac{u-t_0}{s}\Big|+\Big|\frac{v-t_1}{s}\Big|\Big)^{\alpha'}\Big) \right. \\
&\quad \left. \Big(\frac{1}{1+\big(|(x-u)/s|+|(y-v)/s|\big)^m}\Big)\, dudv \right)\frac{ds}{s^2}.
\end{aligned}$$

Since $2^{j+1}-2^j = 2^j$ there exists an constant K such that;

$$\begin{aligned}
|\Delta_j(x,y)| \le K(2^{j+1}-2^j)\iint_{-\infty}^{\infty} 2^{j\alpha}\Big(1+\Big(\Big|\frac{u-t_0}{2^j}\Big|+\Big|\frac{v-t_1}{2^j}\Big|\Big)^{\alpha'}\Big) \\
\Big(\frac{1}{1+\big(|(x-u)/2^j|+|(y-v)/2^j|\big)^m}\Big)\,\frac{dudv}{2^{2j}}.
\end{aligned}$$

With the change of variable $u' = 2^{-j}(u-t_0)$ and $v' = 2^{-j}(v-t_1)$, and since $\big(|u-t_0|+|v-t_1|\big)^{\alpha'} \le \big(|u-x|+|x-t_0|+|v-y|+|y-t_1|\big)^{\alpha'} \le$

$2^{\alpha'}\Big(\big(|u-x|+|v-y|\big)^{\alpha'}+\big(|x-t_0|+|y-t_1|\big)^{\alpha'}\Big)$ it follows that;

$$\begin{aligned}|\Delta_j(x,y)| &\le K2^{\alpha j}\iint_{-\infty}^{\infty}\frac{1+(|u'|+|v'|)^{\alpha'}+\Big(|(t_0-x)/2^j|+|(t_1-y)/2^j|\Big)^{\alpha'}}{1+(|u'|+|v'|)^m}\,du'dv'\\ &\le K2^{\alpha j}\Bigg(\iint_{-\infty}^{\infty}\frac{1+(|u'|+|v'|)^{\alpha'}}{1+(|u'|+|v'|)^m}\,du'dv'\\ &\quad+\Big(|(t_0-x)/2^j|+|(t_1-y)/2^j|\Big)^{\alpha'}\iint_{-\infty}^{\infty}\frac{1}{1+(|u'|+|v'|)^m}\,du'dv'\Bigg).\end{aligned}$$

By choosing $m=\alpha'+2$, both integrals are convergent. Therefore there exists an constant K such that;

$$|\Delta_j(x,y)|\le K2^{\alpha j}\left(1+\Big(\Big|\frac{t_0-x}{2^j}\Big|+\Big|\frac{t_1-y}{2^j}\Big|\Big)^{\alpha'}\right).$$

Let $f^{(x)}(x,y)$ and $f^{(y)}(x,y)$ denote the partial-derivative of f with respect to the variable x and y. Since $f(g(x),h(y))^{(x)}=g^{(x)}(x)f^{(x)}(u,y)\big|_{u=g(x)}$, and Eq.(3.3.1) holds for any partial derivative of ψ^x and ψ^y, a similar computation shows that there exists constants K_1 and K_2 such that;

$$|\Delta^{(x)}(x,y)|\le K_1 2^{(\alpha-1)j}\left(1+\Big(\Big|\frac{t_0-x}{2^j}\Big|+\Big|\frac{t_1-y}{2^j}\Big|\Big)^{\alpha'}\right)$$

$$|\Delta^{(y)}(x,y)|\le K_2 2^{(\alpha-1)j}\left(1+\Big(\Big|\frac{t_0-x}{2^j}\Big|+\Big|\frac{t_1-y}{2^j}\Big|\Big)^{\alpha'}\right).$$

In particular for $(x,y)=(t_0,t_1)$, it follows that;

$$|\Delta_j(t_0,t_1)|\le K'2^{\alpha j}.$$

Since $\alpha>0$ this guarantees fast decay of $|\Delta_j(t_0,t_1)|$ as 2^j goes to zero. By Prop.(B.0.8) it follows that $Mf((u,v),s)\le\|\ f\ \|\ (\|\ \psi^x\ \|+\|\ \psi^y\ \|)$. Then;

$$\begin{aligned}|\Delta(t_0,t_1)| &\le \frac{\|\ f\ \|\ (\|\ \psi^x\ \|+\|\ \psi^y\ \|)}{C_\psi}\\ &\int_0^\infty\iint_{-\infty}^{\infty}|\psi^x(\frac{x-u}{s},\frac{y-v}{s})+\psi^y(\frac{x-u}{s},\frac{y-v}{s})|\,dudv\frac{ds}{s^3}\end{aligned}$$

With a change of variable $su' = x - u$ and $sv' = y - v$ it follows that;

$$\begin{aligned}|\Delta(t_0, t_1)| &\leq K \iint_{-\infty}^{\infty} |\psi^{x)}(u', v') + \psi^{y)}(u', v')| \, du'dv' \int_{2^j}^{2^{j+1}} \frac{ds}{s^2} \\ &\leq K 2^{-(1+j)}.\end{aligned}$$

From this it follows that the coefficients converges for 2^j large.

The computations up to this point have verifies the boundedness of the constant in the beginning of the proof. The next step is to verify that f satisfy the local Lipschitz-condition at (t_0, t_1). By using the sum-expressions of f and C;

$$|f(x, y) - C| = \Big| \sum_{j=-\infty}^{\infty} \Big(\Delta_j(x, y) - \Delta_j(t_0, t_1) \Big) \Big|.$$

The next task is to verify that $|f(x+h, y+h) - C| < A(|h| + |k|)^{\alpha}$, for some constant A and $|h|, |k| < \epsilon_h, \epsilon_k$.

The sum in the latter equation will be divided into two at $2^J \geq |t_0 - x| + |t_1 - y| \geq 2^{J-1}$. Note that for $j \leq J$ the sum, everything is continuously differentiable. For $j \geq J$ it follows by the Taylor-error formula that;

$$\begin{aligned} I &= \Big| \sum_{j=J}^{\infty} \Big(\Delta_j(x, y) - \Delta_j(t_0, t_1) \Big) \Big| \\ &\leq | \sum_{j=-\infty}^{\infty} \Big((x - t_0) \sup_{|h|<|x|} \Delta_j^{(x)}(h, y) + (y - t_1) \sup_{|k|<|x|} \Delta_j^{(y)}(x, k) \Big) | \\ &\leq K \big(|(x - t_0)| + |(y - t_1|) \big) \sum_{j=J}^{\infty} \frac{2^{\alpha j}}{2^j} \Big(1 + \frac{(|t_0 - x| + |t_1 - y|)^{\alpha'}}{2^{\alpha' j}} \Big). \end{aligned}$$

By using the assumption it follows that;

$$\begin{aligned}
I &= K_1\big(|(x-t_0)|+|(y-t_1|\big)\sum_{j=J}^{\infty}\frac{2^{\alpha j}}{2^j}\Big(1+\frac{(|t_0-x|+|t_1-y|)^{\alpha'}}{2^{\alpha' j}}\Big)\\
&\le K_2\big(|(x-t_0)|+|(y-t_1|\big)\sum_{j=J}^{\infty}\frac{2^{\alpha j}}{2^j}\Big(\frac{(|t_0-x|+|t_1-y|)^{\alpha'}}{2^{\alpha' j}}\Big)\\
&\le K_3 2^J\sum_{s=0}^{\infty}\frac{2^{\alpha J}2^{\alpha s}}{2^J 2^s}\Big(\frac{(2^J)^{\alpha'}}{2^{\alpha' s}2^{\alpha' J}}\Big)\le 2^{\alpha}K_4 2^{\alpha(J-1)}\sum_{s=0}^{\infty}\frac{2^{\alpha s}}{2^{s(1+\alpha')}}\\
&\le K_5\Big(|t_0-x|+|t_1-y|\Big)^{\alpha}
\end{aligned}$$

On the other hand it follows that for small j;

$$\begin{aligned}
II &= \Big|\sum_{j=-\infty}^{J-1}\Big(\Delta_j(x,y)-\Delta_j(t_0,t_1)\Big)\Big|\\
&\le K_1\Big|\sum_{j=-\infty}^{J-1}\Big\{2^{\alpha j}\Big(1+\Big(\Big|\frac{t_0-x}{2^j}\Big|+\Big|\frac{t_1-y}{2^j}\Big|\Big)^{\alpha'}\Big)-2^{\alpha j}\Big\}\Big|\\
&\le K_2\sum_{j=-\infty}^{J-1}\Big\{2^{\alpha j}+2^{\alpha j-\alpha' j}\Big(|t_0-x|+|t_1-y|\Big)^{\alpha'}+2^{\alpha j}\Big\}\\
&\le K_3\Big(2^{\alpha J}+2^{(\alpha-\alpha')J}\Big(2^J\Big)^{\alpha'}+2^{\alpha J}\Big)\\
&\le K_4 2^{\alpha}\Big(2^{\alpha(J-1)}+2^{\alpha(J-1)}+2^{\alpha(J-1)}\Big)\\
&\le K_5(|x-t_0|+|y-t_1|)^{\alpha}.
\end{aligned}$$

From this it follows that;

$$|f(x,y)-f(t_0,t_1)|\le I+II\le A(|x-t_0|+|y-t_1|)^{\alpha},$$

i.e. $f(x,y)$ is locally Lipschitz-α at (t_0,t_1).

□

The unfortunate thing after used so much time on a proof is that the result will not be used in practice. Assume the wavelet has compact support in both

x- and y-direction. Then by a similar argument as in 1-D, one may verify that if the spatial position of an modulus-maximum changes little across scales one can estimate the local Lipschitz-regularity by the following equation;

$$\alpha = \frac{\log_2(|Mf((u,v),s_1)|) - \log_2(|Mf((u',v'),s_0)|)}{\log_2(s_1) - \log_2(s_0)} - 1.$$

(u,v) and (u',v') are two related modulus-maximum at scales s_1 and s_0.

In Ch.(2), Th.(2.5.1) reduced the number of computations in order to find the Lipschitz-regularity of a function, by proving that it suffices to consider the decay of the wavelet-transform at the modulus-maximum. It will not be attempted to verify an equivalent statement in 2-D of two reasons. First, an eventual proof would be very long. Second, this statement would have no practical applications in the discussion of 2-D edge-detectors, so the reward for investigating such a proof would be little. In this thesis be assumed that if $Mf(u,s)$ has no modulus-maximum in $]a,b[\times]c,d[$ as $s \to 0$, then f is globally Lipschitz-1 at $]a+\epsilon, b-\epsilon[\times]c+\epsilon, d-\epsilon[$. Another uncertainty in 2-D is how/if a local maxima of $Mf(u,s)$ converges towards finer scale. In Ch.(5) assumptions is used in order to relate modulus-maximum across scales.

Chapter 4

Noise.

In this chapter some concepts concerning the presence of noise in a signal will be discussed.

One component which typically makes edge-detection difficult is the existence of noise in signals. Detecting edges in signals without noise is typically an easy task. In this case there are typically only minor differences between the performance of different edge-detectors. The problem of detecting edges in noisy signals is how one can separate modulus-maximum caused by noise, texture and less important structures in the signal, from the modulus-maximum caused by important edges. This is further discussed in Ch.(6) and Ch.(7).

The tools and concepts discussed in this chapter will serve a two-folded purpose. First, the Wiener-Filtering and the SNR introduced in Sect.(4.2) and Sect.(4.3) will be the fundation of some edge-detectors presented later in this thesis. The second purpose of the discussion is to support some observations which will be made about edge-detectors and the behaviour of modulus-maximum across scales.

The chapter begins with discussing *Gaussian White Noise*. The main-result of this discussion is that the continuous wavelet-transform preserves information about the noise in a signal. Secondly, the uncertainty-principle of noisy signals will be discussed. The uncertainty-principle implies that it is difficult to obtain both good detection and good localization of the edges. The last section will discuss how one may estimate the noise in a signal. This estimate will be used in Sect.(6.1) and Sect.(7.1.1) to construct an edge-detector.

4.1 Additive Gaussian White Noise.

This section presents *additive Gaussian white noise* and its response with respect to the CWT.

The presence of noise is what makes applied edge-detection difficult and interesting. When constructing edge-detectors for signals which contain noise, one possible approach is to utilize differences between noise and the signal itself. Noise is a random process, and may be considered as the realisation of some probability-distribution. It makes little sense to discuss noise at a single point, but it makes sense to discuss the overall influence of noise in the signal.

It will be proved that the CWT preserves information about the distribution of noise in a signal. This relation will be used in Sect.(4.3) to estimate the noise in the image in Fig.(1.4).

The section begins with a description of additive Gaussian white noise, before establishing the relation between noise and the wavelet-transform.

In this thesis the noise will be assumed to be *additive noise*. This mean that for a signal f, it is assumed to exist a noise-free signal f_0 and a random process n such that;

$$f(t) = f_0(t) + n(t).$$

In this thesis the noise is assumed to be *White Gaussian Noise*.

Definition 4.1.1 (Gaussian White Noise.) *Gaussian white noise of variance σ^2 denoted n_σ is a random process with zero-mean such that;*

$$E[n_\sigma(t_1)n_\sigma(t_2)] = \sigma^2\delta(t_1 - t_2).$$

$E[\cdot]$ denotes the expected-value.

□

The definition implies that the noise-coefficients have zero-mean, are distributed as a Gaussian with variance σ^2 and are uncorrelated. That the noise-coefficients are uncorrelated mean that the influence of noise at a single point is independent of the noise in the neighbourhood of the point.

The next statement will establish a relation between noise and its response to the CWT.

Statement 4.1.1 *Assume $n_\sigma(t)$ is Gaussian white noise with variance σ^2. Then;*

$$E[Wn_\sigma(u,s)^2] = \sigma^2 \parallel \psi_s \parallel_2^2 .$$

Proof: Assume the wavelet-transform is computed with respect to the wavelet ψ_s. By Def.(4.1.1) it follows that;

$$\begin{aligned} E[Wn_\sigma(u,s)^2] &= E\Big[\int_{-\infty}^{\infty} n_\sigma(t)\psi_s(t-u)\,dt \int_{-\infty}^{\infty} n_\sigma(\tau)\psi_s(\tau-u)\,d\tau\Big] \\ &= E\Big[\int_{-\infty}^{\infty} n_\sigma(t)n_\sigma(\tau)\psi_s(t-u)\psi_s(\tau-u)\,dtd\tau\Big] \\ &= \int_{-\infty}^{\infty} E[n_\sigma(t)n_\sigma(\tau)]\psi_s(t-u)\psi_s(\tau-u)\,dtd\tau \\ &= \sigma^2 \int_{-\infty}^{\infty} \psi_s(t-u)\psi_s(t-u)\,dt = \sigma^2 \parallel \psi_s \parallel_2^2 . \end{aligned}$$

□

The relation in Stat.(4.1.1) will be used in Sect.(4.3) to estimate σ^2.

4.2 The Uncertainty Principle in noisy signals.

This section will discuss the uncertainty-principle of edge-detection in noisy signals, introduced by Canny in [3].

Two factors play a critical role when detecting the edges in a signal. The first target is that one has to detect all the important edges and as few unwanted edges as possible. The second target it that the position of the detected modulus-maximum must be as close as possible to the position of the edges. This section introduces two criteria used to quantify detection- and localization-performance of an edge-detector, the *Signal-to-Noise Ratio* (SNR) and *Location* (LOC).

Consider Fig.(2.7) illustrating the modulus-maximum at number of scales of the signal in Fig.(2.7a). At large scales one may observe that the amplitude of the modulus-maximum corresponding to noise is small compared to the modulus-maximum corresponding to the important edges in the signal. This

makes it easier to distinguish between modulus-maximum corresponding to noise and important edges by using the amplitude of the modulus-maximum. At small scales the spatial position of the modulus-maximum improves, but it is more difficult to distinguish between modulus-maximum corresponding to noise and important edges. In this section quantifications of these observations will be discussed. This will establish a relation between detection- and localization-performance, resulting in the uncertainty-principle for noisy-signals.

It will be derived that for step-edges the detection and localization performance are inversely related. In practice this implies that there exists no single scale such that the output of the CWT gives both optimal detection and optimal localization of the edges. This will in particular be the case for signals where the influence of noise is large, e.g. US-images.

The section begins with introducing the quantities which will be used to quantify detection-and localization-performance of the CWT. It will be derived that detection- and localization-performance is inversely related. Consequences of this relation will be discussed.

Definition 4.2.1 (Signal-to Noise-Ratio. [3]) *Assume* $f(t) = f_0(t) + n_\sigma(t)$ *where* $n_\sigma(t)$ *is additive Gaussian noise. The output signal-to-noise ratio of* $Wf(u,s)$ *at a point* u_0 *and scale* s_0 *is defined as;*

$$SNR(u_0, s_0) = \frac{|Wf(u_0, s_0)|}{\sigma \parallel \psi \parallel_2},$$

where σ^2 *is variance of the noise.*

□

The underlying idea behind the concept is simple. The numerator is the absolute-value of the CWT. The denominator can be considered as the expected value of the CWT for the noise only. This was verified by Stat.(4.1.1). If SNR at a modulus-maximum is small, there is a big possibility that the modulus-maximum is caused by noise. If the value is large it is a small possibility that the modulus-maximum is caused by noise.

Definition 4.2.2 (Localization Criteria. [3]) *Assume* $f(t) = f_0(t)+n_\sigma(t)$ *where* $n_\sigma(t)$ *is additive Gaussian noise. The Localization-ratio of* $Wf(u,s)$ *at the point* u_0 *and scale* s_0 *is defined as;*

$$LOC(u_0, s_0) = \frac{|\int_{-\infty}^{\infty} f'(t)\psi'_{s_0}(t-u_0)\,dt|}{\sigma\sqrt{\int_{-\infty}^{\infty}(\psi'_{s_0})^2(t)\,dt}}.$$

σ^2 *is the variance of the noise.*

□

The localization of an edge improves as LOC increases [3]. I.e. the target is to get as high value of LOC as possible to obtain as good localization as possible.

The next step is to consider the behaviour of SNR and LOC for a step-edge with amplitude A at 0. Assume that $f(t) = Au(t) + n_\sigma(t)$ where $n_\sigma(t)$ is noise with variance σ^2. For $s = 1$ the expressions for SNR and LOC are;

$$SNR(0,1) = \frac{A|\int_0^{\infty}\psi(t)\,dt|}{\sigma \parallel \psi(t) \parallel} = \frac{A}{\sigma}\Sigma(\psi), \quad \Sigma(\psi) = \frac{|\int_0^{\infty}\psi(t)\,dt|}{\parallel \psi(t) \parallel}$$
$$LOC(0,1) = \frac{A|\psi'(0)|}{\sigma\sqrt{\int_{-\infty}^{\infty}\psi'(t)^2\,dt}} = \frac{A}{\sigma}\Lambda(\psi'), \quad \Lambda(\psi') = \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty}\psi'(t)^2\,dt}}.$$

For $s \neq 1$ the corresponding expressions for $\Lambda(\psi'_s)$ and $\Sigma(\psi_s)$ are;

$$\Sigma(\psi_s) = \frac{s^{-1/2}|\int_0^{\infty}\psi(t/s)\,dt|}{\parallel \psi(t) \parallel} = \frac{s^{1/2}|\int_0^{\infty}\psi(y)\,dy|}{\parallel \psi(t) \parallel} = \sqrt{s}\Sigma(\psi)$$
$$\Lambda(\psi'_s) = \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty}\psi'(t/s)^2\,dt}} = \frac{|\psi'(0)|}{s^{1/2}\sqrt{\int_{-\infty}^{\infty}\psi'(y)^2\,dy}} = \frac{1}{\sqrt{s}}\Lambda(\psi').$$

From these calculations it follows that the values of SNR and LOC are inversely related with respect to the scale. This relation implies that the detection performance decrease at fine scales and the localization performance decrease at large scales. This support the observations made of the signal in Fig.(2.7). At fine scales the value of SNR is low; causing difficulties in separating modulus-maximum caused by noise and edges. At coarse scale the value of LOC is low; causing bad localization of the modulus-maximum.

Another observation which may be observed from the calculations is that SNR and LOC can not be large simultaneously. Note that $SNR(u,s) \cdot LOC(u,s)$ equals an constant for all $s \in]0,\infty]$. This implies (in theory) that there exists no single scale s such that the edges detected by $Wf(u,s)$ are optimal with respect to both localization and detection-performance. In practice the question of optimality depends on the strength of noise in the signal. If the influence of noise is little there is a big possibility that one can have single-scale edge-detectors which will produce an "almost" optimal output, at least from an visual viewpoint. If the influence of noise is large, there is a small possibility that one can not achieve both good localization and good detection at a single scale. US-images are typically heavily influenced by noise. This implies that there will probably not exist a scale such that one detects all the important edges in the image and that the localization of these edges is good.

This does not imply that one will stop the investigation of edge-detectors for US-images at this point. In Ch.(5) a method which enables to bypass the relation between SNR and LOC will be discussed. This is possible by relating modulus-maximum across scales. This will enable us to find the important edges at coarse scales where SNR are high, and trace these modulus-maximum towards finer scales where LOC is high. This will enable good localization- *and* good detection-performance.

4.3 Wiener-filtering.

This section presents a method which estimates the noise in a signal by using a *Wiener-filter*.

In the previous section SNR and LOC was introduced as a quantity of the detection- and localization-performance of an edge-detector. Both depended on the variance of the noise. In Sect.(6.1) and Sect.(7.1.1) an edge-detector will be presented based on the noise-estimate and SNR. The Canny edge-detector uses the probability distribution of noise to separate the modulus-maximum corresponding to noise from the modulus-maximum corresponding to the important edges. How this edge-detector performs is critically dependent on the estimate of the noise.

The section discusses the method Canny used in [3] to find the Wiener-filter, and how he estimated the noise in signals. At the end this method is applied to the signal in Fig.(2.7) and image in Fig.(1.4).

Assume that the noise is added Gaussian white noise. The coefficients of the wavelet-transform of the noise should be distributed as a Gaussian with variance $\sigma^2 \parallel \psi_s \parallel_2^2$. This section will explain how $Wn_\sigma(u,s)$ can be obtained from $Wf(u,s)$ by using a Wiener-filter. The Wiener-filter tries to find a filter F which separates one of the components of a two-component function. Assume that the function $f(t) = f_0(t) + n_\sigma(t)$, where $n_\sigma(t)$ is noise. Linearity of the wavelet-transform implies that $Wf(u,s) = Wf_0(u,s) + Wn_\sigma(u,s)$. The component we wish to isolate is $Wn_\sigma(u,s)$, from which σ^2 can be estimated by Stat.(4.1.1).

The Wiener-filter uses the auto-correlation of $Wn_\sigma(u,s)$ and $Wf_0(u,s)$, and is given by [3];

$$\begin{aligned} R_1(\tau) &= \int_{-\infty}^{\infty} Wf_0(u,s) Wf_0(u+\tau,s)\, du \\ R_2(\tau) &= \int_{-\infty}^{\infty} Wn_\sigma(u,s) Wn_\sigma(u+\tau,s)\, du \\ R_3(\tau) &= \int_{-\infty}^{\infty} Wf_0(u,s) Wn_\sigma(u+\tau,s)\, du. \end{aligned}$$

The noise and signal is assumed to be independent so $R_3 = 0$. In Ex.(2.1.2) it was verified that $W(k\delta)(u,s) = k\tilde{\psi}_s(u)$, where k is the amplitude. It follows that;

$$\begin{aligned} R_2(\tau) &= k^2 \int_{-\infty}^{\infty} \tilde{\psi}_s(u)\tilde{\psi}_s(u+\tau)\, du = k^2 \int_{-\infty}^{\infty} \psi_s(u)\psi_s(u+\tau)\, du \\ &= \frac{-k^2}{4\sqrt{\pi}}\Big(\frac{\tau^2}{2s^2} - 1\Big) e^{-\frac{\tau^2}{4s^2}}. \end{aligned}$$

As verified by Ex.(2.1.2), $Wf_0(u,s) = As\theta_s(u)$, assumed that f_0 has a step-edge of amplitude A at 0. The auto-correlation is given by;

$$\begin{aligned} R_1(\tau) &= A^2 s^2 \int_{-\infty}^{\infty} \theta_s(u)\theta_s(u+\tau)\, du \\ &= \frac{A^2 s^2}{2\sqrt{\pi}} e^{-\frac{\tau^2}{4s^2}}. \end{aligned}$$

The filter F is implicitly defined [3] as the filter satisfying;

$$R_2(\tau) = (R_1 + R_2) * F(\tau). \tag{4.3.1}$$

If the amplitude of the step-edge is large compared with the amplitude of the noise, one can assume that $R_1 + R_2 \approx R_1$, i.e. the filter F is defined as the filter which satisfies;

$$k_1\Big(\frac{\tau^2}{2s^2} - 1\Big)e^{-\frac{\tau^2}{4s^2}} = F * k_2 e^{-\frac{\tau^2}{4s^2}},$$

for some constants k_1 and k_2. A filter of the form;

$$F(\tau) = \Big(\frac{\tau^2}{s^2} - 1\Big)e^{-\frac{\tau^2}{2s^2}}$$

can be used as an approximation [3] for F. The filter F was defined to satisfy;

$$Wf(\cdot, s) * F(u) = (Wf_0(\cdot, s) + Wn_\sigma(\cdot, s)) * F(u) = Wn_\sigma(\cdot, s) * F(u),$$

i.e. the noise-free signal is removed.

The next step is to estimate σ^2 from $Wn(\cdot, s) * F(u)$. Stat.(4.1.1) verified that $E[Wn(u, s)^2] = \sigma^2 \parallel \psi_s \parallel_2^2$. Assume that the wavelet-transform is calculated for $u \in \Omega$, $\Omega \subseteq \mathbb{N}$. Then the variance of the values of $Wn_\sigma(\cdot, s)$ can be estimated by;

$$E[Wn_\sigma(\cdot, s)^2] = \frac{1}{|\Omega'|}\sum_{i \in \Omega'} Wn_\sigma(u_i, s)^2.$$

Ω' is a subset of Ω. The next paragraph will determine suitable subsets Ω' which are can be used to estimate the $E[Wn_\sigma(\cdot, s)^2]$.

Canny discussed in [3] two approaches for estimating the noise. The simplest is to estimate σ of all the values, $\Omega' = \Omega$. Canny argued in [3] that this may be heavily affected by modulus-maximum corresponding to edges with large amplitude. He rather used some fixed-percent of the modulus-maximum of smallest amplitude to compute the average amplitude, i.e. Ω' equals the points in Ω of smallest value. Modulus-maximum caused by noise, occurs frequently with a Gaussian distribution, while modulus-maximum corresponding to edges occur infrequently and with larger values. Cannys method for computing σ^2, was to compute the average of the 80% lowest values. Tab.(4.1) displays the estimate values of σ^2 using different percentage values. Compared with the example in Sect.(2.7) one may observe that using 95% of the lowest value to estimate σ results in a noise-estimate which is close to the "guessed" estimate in Sect.(2.7).

	100 %	99 %	95 %	80 %
1-D	11	10	8.4	5.4
2-D	11	10.2	8.4	5.5

Table 4.1: Table of noise-estimates for the 1-D signal in Fig.(2.13) and the US-image in Fig.(1.4). Different percentage-values have been used to calculate the noise-estimates. These values will be used in Ch.(6) and Ch.(7).

Since noise at a point is assumed independent on the values in a neighbourhood of the point, a similar approach has been used to estimate the noise of the US-image in Fig.(1.4). Recall that the signal in Fig.(2.13) is a horizontal ray of the US-image. To estimate the noise in the image, the writer has used the above approach to every horizontal line of the US-image. One may observe from Tab.(4.1) that there is only a minor difference between the noise-estimates corresponding to the signal in Fig.(2.13) and the image of Fig.(1.4).

Part II

Edge-Detectors.

Chapter 5

The Maxima-Tree.

In this chapter it will be investigated how one may connect modulus-maximum across scales in 1-D and 2-D.

Connecting modulus-maximum across scales is necessary in order to take advantage of the zooming-property of the wavelet-transform. It is desirable to know which modulus-maximum at fine scales correspond to which modulus-maximum at coarse scales. In Sect.(4.2) it was discussed that localization- and detection-performance of edge-detectors are inversely related. At coarse scales it is easier to distinguish modulus-maximum corresponding to important edges from the less important. At fine scales the spatial position of the modulus-maximum is good with respect to the position of the edges in the signal. The target is to connect modulus-maximum across scales which correspond to similar features in the signal. This relation makes possible edge-detectors with both good localization- and detection-performance. Another motivation for relating modulus-maximum across scales is in order to estimate the local Lipschitz-regularity of the modulus-maxima. Th.(2.5.1) verified that it was sufficient to estimate the decay of the wavelet-transform at the modulus-maximum only. To do this one has to know which modulus-maximum should be computed versus each other.

There will always be some uncertainty when connecting modulus-maximum across scales. There is a huge diversity of functions and signals. It is a challenge to construct a connecting algorithm suitable for all functions, at least if the computer-efforts should be kept to a minimum. How fast and how much the signal oscillates and which kind of edges are in the function, are some of the factors one have to take into account when constructing an connection-algorithm. The main focus of the discussion is to connect modulus-maximum

across scales with respect to the signals which are investigated in this thesis.

A surprising bonus of the connection procedure is that the maxima-tree itself may be used as a simple edge-detector. The maxima-tree will not be discussed as an edge-detector, but it will be illustrated how one may use the maxima-tree to detect the important edges in the signal. In Ch.(6) and Ch.(7) edge-detectors taking advantage of the maxima-tree will be presented. It will be discussed that these edge-detectors are superior to edge-detectors which only uses the modulus-maximum at a single scale with respect to the type of signals analysed in this thesis.

In Sect.(5.1) a 1-D connection-procedure will be presented. How this algorithm is derived, and how it performs will be discussed. In 2-D there are (at least) two approaches for how modulus-maximum can be connected across scales. One may consider the convergence of single modulus-maximum or the convergence of line-segments (collection of modulus-maximum). How this can be done will be discussed in Sect.(5.2). One should note that the 2-D algorithms are not optimal, i.e. false connections may occur. In particular the discussion of convergence of line-segments across scale is ad hoc. The aim of the section is to describe how the problems of convergence have been tried solved with respect to the image in Fig.(1.4) and similar US-images. This is necessary in order to discuss and illustrate the effect to 2-D multi-scale edge-detectors.

5.1 1-D Maxima-Tree.

In this section an algorithm creating the 1-D *modulus-maximum tree* or *maxima-tree* will be presented. The motivation behind this construction and how it is derived will be discussed.

The proposed connection procedure which will be presented requires two steps. The first step is to find the set of modulus-maximum which possibly can be connected across scales. This step based on the *sign* of the modulus-maximum, and is funded on Prop.(2.5.1). The set of possible connections is typically redundant. The second step is to remove the redundancy in the set of possible connections. The first step is based on theoretic facts, i.e. they are valid for all signals. The second step is based on assumptions. One can therefore not expect that the proposed algorithm holds for all signals. It will be discussed when one should suspect problems with the algorithm.

The section begins with a brief account of which information of the

modulus-maximum are available for constructing the maxima-tree. Two properties of the *sign* of the modulus-maximum will be derived. The redundancy of the tree-construction and an algorithm removing the redundancy will be discussed. At the end the proposed method is briefly compared with a few alternative methods for connecting modulus-maximum across scales.

In Sect.(2.3) the wavelet-based trivial edge-detector was defined. The output of the trivial edge-detector is all the modulus-maximum at the scales the wavelet-transform have been computed. Assume the wavelet-transform is computed for $s \in \{s_1, ..., s_J\}$ with $s_{i-1} < s_i$. The following information is available from the output;

- The *position* $\{(m_i^j, s_j)\}$ of the modulus-maximum, where $i \in \{1, ..., N_j\}$ and $s_j \in \{s_1, ...s_J\}$. N_j is the number of modulus-maximum at scale s_j.
- The *amplitude* $a_i^j = |Wf(m_i^j, s_j)|$ of the modulus-maximum for $i \in \{1, ..., N_j\}$ and $s_j \in \{s_1, ...s_J\}$.
- The *sign* of the modulus-maximum.

By Th.(2.5.2) it follows that if $Wf(u, s)$ is computed for a dense set of scales, one may easily determine which modulus-maximum should be connected across scales. In practice one can only compute the CWT for a discrete set of scales. Computing the CWT is typically one of the more time-consuming operations in edge-detection algorithms. The less scales for which one has to compute the wavelet-transform, the less computer-effort is required. In order to determine which modulus-maximum are connected across scales, one need to exploit some additional information about the modulus-maximum. The following two properties help to decide which modulus-maximum can be a member of the same maxima-line, see Def.(2.5.1).

Property 5.1.1 *Assume ψ^1 is the wavelet given by Eq.(2.1.4) and that the wavelet-transform has been computed with respect to ψ^1. The sign of a modulus-maximum can not change along a maxima-line.*

Proof: Assume (m_0, s_0) and (m_1, s_1) are two modulus-maximum of $Wf(u, s)$, and assume that they are on the same maxima-line. Assume $0 < s_0 < s_1$ and $sign(a_0) \neq sign(a_1)$, where a_0, a_1 are the amplitude of the modulus-maximum. By Th.(2.5.2) the maxima-line is connected, and Prop.(2.5.1)

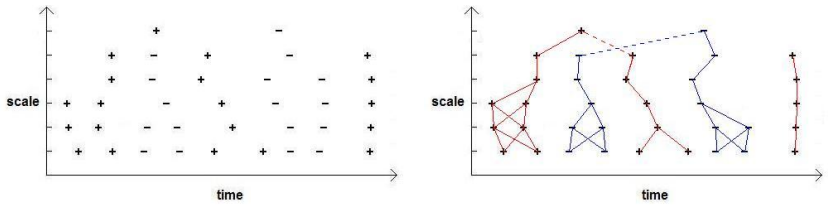
implies that there exists a modulus-maximum (m, s) with $s_0 < s < s_1$ and $a = 0$. If $a = 0$, the point can not be a modulus-maximum, contradicting that (m, s) is a modulus-maximum.

□

Property 5.1.2 *Assume ψ^1 is the wavelet given by Eq.(2.1.4) and that the wavelet-transform has been computed with respect to ψ^1. Two maxima-lines of opposite sign can not cross.*

Proof: Assume (m_0, s_0) and (m_1, s_1) are two related modulus-maximum of positive sign, and (y_0, s_0) and (y_1, s_1) are two related modulus-maximum of negative sign. Assume that $m_0 > y_0$ and $m_1 < y_1$. Prop.(2.5.1) and Th.(2.5.2) imply that there exists a $s_0 < s < s_1$ such that $(m, s) = (y, s)$, where m is connected to the maxima-line of m_0 and m_1 and y is connected to the maxima-line of y_0 and y_1. Denote by a_m and a_y the amplitude of the modulus-maximum at scale s. Then $a_s = a_y \Rightarrow a_s = a_y = 0$ contradicting that they are modulus-maximum.

□



(a) (b)

Figure 5.1: (a) The "+" corresponds to modulus-maximum with positive amplitude. The "-" corresponds to modulus-maximum with negative amplitude. (b) The relation between the modulus-maximum by using Prop.(5.1.1) and Prop.(5.1.2). The dotted lines correspond to possible maxima-lines where the two properties can not determine which are related.

Fig.(5.1) illustrates how the two properties establish the set of possible connections of modulus-maximum across scales. One may observe that if two (or several) maxima-lines of similar *sign* are not separated by a maxima-line

of opposite *sign*, the properties can not tell which modulus-maximum are related in a unique way. This redundancy will be attempted removed in the next step of the algorithm. An error which may be difficult to remove is illustrated by the dotted lines in Fig.(5.1b). None or one of the dotted lines can be solid, but not the both. This may or may not be fixed by the next step of the algorithm. One can determine which, if any, of the maxima-lines should be connected by computing the CWT for some additional scales in between the scales which causes difficulties.

Before presenting how one may remove the redundancy in the set of possible connections achieved by using Prop.(5.1.1) and Prop.(5.1.2), the behaviour of the maxima-line will be investigated. This discussion is restricted to its simplest form; the function $f(t) = u(t) + Au(t-1)$ for $A > 1$, where $u(t)$ is the Heaviside. I.e. f has step-edges at 0 and 1, and f changes more at 1 than at 0. The discussion will be done with respect to the wavelet ψ^1 in Eq.(2.1.4). By Ex.(2.1.2) it follows that;

$$\begin{aligned} Wf(\tau,s) &= \int_0^\infty \frac{1}{\sqrt{s}}\psi_s^1(t-\tau)\,dt + \int_1^\infty \frac{1}{\sqrt{s}}\psi_s^1(t-\tau)\,dt \\ &= s\theta_s(\tau) + As\theta_s(\tau-1). \end{aligned}$$

By symmetry of the Gaussian, all modulus-maximum of $Wf(\tau,s)$ are within $]0,1[$ for all $s \in \mathbb{R}_+$. The maxima-lines are found by solving the following equation for $(\tau,s) \in \mathbb{R}_+ \times]0,1[$;

$$\begin{aligned} \frac{dWf(\tau,s)}{du} &= -\psi_s^1(\tau) - A\psi_s^1(\tau-1) = 0 \\ &\Downarrow \\ 0 &= \tau e^{-\frac{\tau^2}{2s^2}} + A(\tau-1)e^{-\frac{(\tau-1)^2}{2s^2}} \\ &\Downarrow \\ \tau e^{-\frac{\tau^2}{2s^2}} &= A(1-\tau)e^{-\frac{(\tau-1)^2}{2s^2}} \\ &\Downarrow \\ \ln(\tau) - \frac{\tau^2}{2s^2} &= \ln(A) + \ln(1-\tau) - \frac{(\tau-1)^2}{2s^2} \\ &\Downarrow \\ g(\tau) := \ln(\frac{\tau}{1-\tau}) &= \ln(A) + \frac{1}{s^2}(\tau - 1/2) =: f(\tau). \end{aligned}$$

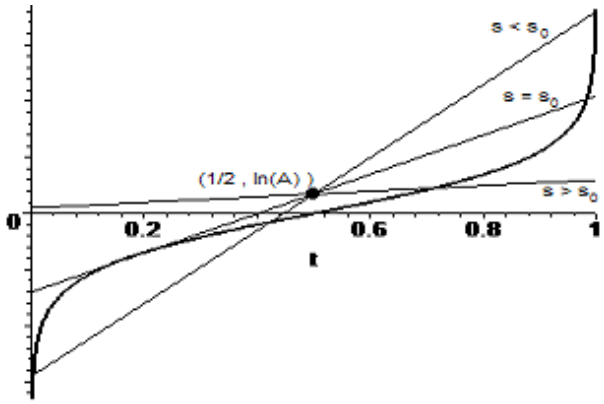


Figure 5.2: The graphs of $f(\tau)$(thin lines) and $g(\tau)$(thick line) for some $A > 1$. The line marked $s > s_0$ corresponds to $f(\tau)$ at coarse scales before the maxima-line which converges to the small edge appears. The line marked $s = s_0$ corresponds to $f(\tau)$ at the scale the second modulus-maxima appear, and the line marked $s < s_0$ corresponds to $f(\tau)$ at small scales.

Note that all factors in the 3.rd equality are strictly positive, and note from the last equality that the scale determine the slope of $f(\tau)$. Consider Fig.(5.2) displaying the values of $f(\tau)$ and $g(\tau)$ for $u \in]0,1[$. The points of intersections between $g(\tau)$ and $f(\tau)$ are the candidate modulus-maximum of $Wf(\tau, s)$. From this the following consequences emerge;

- As $s \to \infty$ there is only one modulus-maximum. This is positioned at $\tau = \frac{A}{1+A}$.
- The maxima-line originating at $\lim_{s\to\infty}(\frac{A}{1+A}, s)$, converges towards the step-edge of maximal strength.
- The maxima-line at coarse scales converges towards both step-edges *if and only if* $A = 1$.
- If $A > 1$ there exists an $s_0 > 0$ such that for $s > s_0$ there is only one solution, $s = s_0$ two solutions and for $s < s_0$ three solutions of the equation $g(\tau) = f(\tau)$ for $\tau \in]0,1[$. The intersection to the left and right converge towards resp. 0 and 1 as $s \to 0$. The intersection in the center converges towards 1/2. The right and left intersection correspond to

modulus-maximum. The middle intersection corresponds to a local minimum.

- s_0 depends on A. The larger A the smaller value of s_0, i.e. the finer scale before the small step-edge will be detected. Generally one can say that the larger ratio between two step-edges, the smaller scale before the small step-edge is detected.

- The larger value of A, the larger minimal *spatial distance* between the maxima-lines converging to 0 and 1.

Note in particular that the scale at which the small edge is detected, does not depend of the strength of the edge only, but of the strength of other edges in its proximity. The term *spatial distance* of two modulus-maximum (x, s_1) and (y, s_2) will be denoted $d(x, y)$ and is defined as $d(x, y) = |x - y|$. Let $l_a : \mathbb{R}_+ \rightarrow]0, 1[$ denote the maxima-line which converges towards a. $l_a(s)$ is the modulus-maximum at scale s of the maxima-line which converges towards a.

Statement 5.1.1 *Let $f(t) = u(t) + Au(t-1)$ for $A > 1$, where $u(t)$ denotes the Heaviside. Let the wavelet-transform be computed with respect to the wavelet $\psi^1(t)$. For all $\gamma \in]0, 1/2[$ there exists an $A > 1$ such that l_0 first appears at γ, where l_0 is the maxima-line which converge towards 0. The scale s_0 at which l_0 first appears is given by;*

$$s_0 = \sqrt{\gamma - \gamma^2}.$$

The strength A of the step-edge at 1 is given by;

$$A = \frac{\gamma}{1-\gamma} e^{\frac{-1}{2(\gamma^2-\gamma)}} e^{\frac{1}{(\gamma-1)}}.$$

Assume (x_1, s_1) and (x_2, s_2) are two elements of l_0 and (y_1, s_1) and (y_2, s_2) are two elements of l_1. Then;

$$\begin{aligned} d(x_1, x_2) &= |x_1 - x_2| < \gamma \\ d(y_1, y_2) &= |y_1 - y_2| < \gamma \\ d(y_1, x_1) &= |y_1 - x_2| > 1 - 2\gamma. \end{aligned}$$

Proof: Assume $\gamma \in]0, 1/2[$. At the scale which the modulus-maximum first appears, the slope of the function f is equal the slope of g. I.e.

$$f'(\gamma) = g'(\gamma) = -\frac{1}{\gamma^2 - \gamma} > 0$$
$$\Downarrow$$
$$f(t) = -\frac{1}{\gamma^2 - \gamma} t + C$$

Since $f(a) = g(a)$ at the point the maxima-line first appears;

$$f(\gamma) = -\frac{1}{\gamma^2 - \gamma}\gamma + C = \ln\Big(\frac{\gamma}{1-\gamma}\Big) = g(\gamma)$$
$$\Downarrow$$
$$C = \ln\Big(\frac{\gamma}{1-\gamma}\Big) + \frac{\gamma}{\gamma^2 - \gamma}$$

As previously mentioned, the gradient of $f(t)$ determines the scale of which the wavelet-transform is computed. I.e. the scale at which m_0 appears is given by;

$$s_0 = \sqrt{\Big(-\frac{1}{\gamma^2 - \gamma}\Big)^{-1}} = \sqrt{\gamma - \gamma^2}.$$

The value of $f(1/2)$ equals $\ln(A)$. I.e;

$$\ln(A) = f(1/2) = -\frac{1}{2(\gamma^2 - \gamma)} + \ln\Big(\frac{\gamma}{1-\gamma}\Big) + \frac{\gamma}{\gamma^2 - \gamma}$$
$$A = \frac{\gamma}{1-\gamma} e^{-\frac{1}{2(\gamma^2-\gamma)}} e^{\frac{1}{\gamma - 1}}$$

The inequalities concerning the spatial distance between the maxima-lines follow easily.

□

One consequence of this statement is that if the maxima-line converging to the smallest edge exists, say at (τ, s), then the ratio between the two edges is larger than $\frac{\tau}{1-\tau} e^{\frac{-1}{2(\tau^2-\tau)}} e^{\frac{1}{(\tau-1)}}$.

The next step of relating modulus-maximum across scales is to process the set of possible connections established by Prop.(5.1.1) and Prop.(5.1.2).

The target is to establish which of the possible connections "most likely" correspond to the actual maxima-line. Note that in practice there will always be some uncertainty whether the correct connections are made. This uncertainty depends on the signal which is analysed. Consider the model function $f(t) = u(t) + Au(t-1)$ and assume that $A > 1$. By the remarks prior to Stat.(5.1.1) this implies that a maxima-line converges to one and only one modulus-maxima at the finest scale, or equivalently that each modulus-maximum at coarse scale corresponds to one and only one modulus-maximum at the next finer scale. In one case it is easy to establish the correct connections; If a modulus-maximum can be connected with only one modulus-maximum at the next finer scale, Th.(2.5.2) guarantees that these are connected. The obstacle appears if a modulus-maximum can be connected with several modulus-maximum at the next finer scale, or if several modulus-maximum can be connected with several modulus-maximum at the next finer scale. In the following a method which attempts to establish the correct connections will be presented.

Assume $f(t) = u(t) + Au(t-1), A > 1$. The underlying idea of how the modulus-maximum should be connected across scales is motivated by a connection-algorithm found in [7]. This algorithm will be briefly discussed at the end of the section. Assume a modulus-maximum m can be connected with N modulus-maximum m_i at the next finer scale s. The idea is that the modulus-maximum m should be connected with the modulus-maximum m_i which maximises;

$$|Wf(m_i, s) \cdot D(d_i, \beta)|. \tag{5.1.1}$$

$d_i = |m - m_i|$ is the spatial distance between the modulus-maximum at the two scales, and $D : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ is an appropriate function. The next paragraph will investigate if such function $D(t, \beta)$ exists and eventually which properties $D(t, \beta)$ will have to fulfil.

The function $D : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ will be assumed to satisfy $D(a, s) > D(b, s)$ for $|a| < |b|$, and that $D(t, \beta) > 0,\ \forall (t, \beta) \in \mathbb{R} \times \mathbb{R}_+$. This assumption reflects the observation that spatial distance between two elements of the same maxima-line, is less than if one element is a member of another maxima-line. Assume $f(t) = u(t) + Au(t-1), A > 1$, and $Wf(\tau, s)$ is computed at scales $\{s_1, ..., s_N\}$. In the following it will be discussed what $D(t, \beta)$ has to fulfil in order to make the correct connections by using Eq.(5.1.1).

Case 1: Assume that one modulus-maximum can be connected with only

one modulus-maximum at the next finer scale. By Th.(2.5.2) they should connected. Any function $D(t,\beta)$ which satisfies the above assumptions will make the correct connection.

Case 2: Assume that one modulus-maximum m at scale s', can be connected with two modulus-maximum m_1 and m_2 at the next finer scale s. Assume in addition that $Wf(m_1,s) > Wf(m_2,s)$, i.e. $m_1 > m_2$. This is the case between two scales where the second maximaline appears. Assume $d_1 = |m - m_1|$ and $d_2 = |m - m_2|$. By the observations prior to St.(5.1.1) $d_1 < d_2$, and m should be connected with the modulus-maximum at the next finer scale of largest amplitude. Any function $D(t,\beta)$ which satisfies the above assumptions will make the correct connection.

Case3a: Assume that two modulus-maximum m and m' at scale s_i can be connected with two modulus-maximum m_1, m_2 at the next finer scale s_{i-1}. This is the case if two maxima-lines of similar sign are not separated by a maxima-line of opposite sign. Consider the scenario $Wf(m,s_i) > Wf(m',s_i)$, i.e. $m > m'$. We wish to find which of m_1 and m_2 should be connected with m. By the observations prior to St.(5.1.1), m should be connected with the modulus-maximum of largest amplitude, say m_1. Then $Wf(m_1,s_{i-1}) > Wf(m_2,s_{i-1})$. In addition $d_1 = |m - m_1| < |m - m_2| = d_2$. Any $D(t,\beta)$ will establish the correct connection.

Case3b: Assume that two modulus-maximum m, m' at scale s_i can be connected with two modulus-maximum m_1, m_2 at the next finer scale s_{i-1}. This is the case if two maxima-lines of similar sign are not separated by a maximaline of opposite sign. Consider the scenario $Wf(m,s_i) > Wf(m',s_i)$, and we wish to find which of m_1 and m_2 should be connected with m'. By the observations prior to St.(5.1.1), m' should be connected with the modulus-maximum of smallest amplitude, say m_2. In this case $Wf(m_1,s_{i-1}) < Wf(m_2,s_{i-1})$. By the above observations it follows that $d_1 = |m - m_1| > |m - m_2| = d_2$. In this case the choice of function $D(t,\beta)$ is critical. Not any function $D(t,\beta)$ will establish the correct connection.

As previously discussed, the larger A the smaller scale before the maxima-line converging to the small step-edge appears. This implies that the minimal spatial-distance between the two maximalines increases as A increases. If $D(t,\beta)$ decays sufficiently fast with respect to the scale, one might hope that the correct connections are made in also Case(3b). Let $f(t) = u(t) + Au(t-1)$, and $A > 1$. Assume the wavelet-transform is computed with respect to the wavelet ψ^1, and that l_0 and l_1 are the maxima-lines converging to resp 0 and 1. Assume that l_0 and l_1 exists for all $0 < s < s_0$. Let $0 < s_1 < s_2 < s_0$

and let x_1 and x_2 be points of l_0 at scale s_1 and s_2 and y_1 be a point of l_1 at scale s_1.For the model-function $f(t)$ it follows by Stat.(5.1.1) that there exist a $\gamma \in]0, 1/2[$ such that;

$$\begin{aligned} d(x_2, x_1) &< \gamma, \\ d(x_2, y_1) &> 1 - 2\gamma. \end{aligned}$$

$D(t, \beta)$ will establish the correct connection in Case(3b), the following inequality holds;

$$|Wf(x_1, s_1)D(d(x_2, x_1), \beta)| > |Wf(y_1, s_1)D(d(x_2, y_1), \beta)|. \tag{5.1.2}$$

By St.(5.1.1) and monotony of $\theta(t)$ and $D(t, s)$ for $t \in \mathbb{R}_+$ it follows that;

$$\begin{aligned} |Wf(x_1, s_1)D(d(x_2, x_1), \beta)| &\geq |Wf(x_1, s_1)D(\gamma, \beta)| \\ &= s_1 |\Big(\theta_{s_1}(x_1) + A\theta_{s_1}(x_1 - 1)\Big) D(\gamma, \beta)| \\ &\geq s_1 |\Big(\theta_{s_1}(\gamma) + A\theta_{s_1}(1)\Big) D(\gamma, \beta)| \\ |Wf(y_1, s_1)D(d(x_2, y_1), \beta)| &\leq |Wf(y_1, s_1)D(1 - 2\gamma, \beta)| \\ &= s_1 |\Big(\theta_{s_1}(y_1) + A\theta_{s_1}(y_1 - 1)\Big) D(1 - 2\gamma, \beta)| \\ &\leq s_1 |\Big(\theta_{s_1}(1 - \gamma) + A\theta_{s_1}(0)\Big) D(1 - 2\gamma, \beta)|. \end{aligned}$$

So with respect to Eq.(5.1.2) it is sufficient that;

$$|\Big(\theta_{s_1}(\gamma) + A\theta_{s_1}(1)\Big) D(\gamma, \beta)| > |\Big(\theta_{s_1}(1 - \gamma) + A\theta_{s_1}(0)\Big) D(1 - 2\gamma, \beta)| \tag{5.1.3}$$

For the signal $f(t) = u(t) + Au(t - 1)$, all the components of Eq.(5.1.3) are positive so if the following inequalities hold;

$$\begin{aligned} \theta_{s_1}(\gamma)D(\gamma, \beta) &> \theta_{s_1}(1 - \gamma)D(1 - 2\gamma, \beta) \\ \theta_{s_1}(1)D(\gamma, \beta) &> \theta_{s_1}(0)D(1 - 2\gamma, \beta), \end{aligned}$$

then Eq.(5.1.2) holds. A similar argument holds for all pairs of step-edges of similar *sign*. Assume $D(t, \beta)$ is of the form $D(t, \beta) = e^{-t^2/\beta^2}$ for $\beta \in \mathbb{R}_+$. The above inequalities become;

$$\begin{aligned} e^{-\frac{\gamma^2}{2s_1^2}} e^{-\frac{\gamma^2}{\beta^2}} &> e^{-\frac{(1-\gamma)^2}{2s_1^2}} e^{-\frac{(1-2\gamma)^2}{\beta^2}} \\ e^{-\frac{1}{2s_1^2}} e^{-\frac{\gamma^2}{\beta^2}} &> e^{-\frac{(1-2\gamma)^2}{\beta^2}}. \end{aligned}$$

Since $a < b \Rightarrow \ln(a) < \ln(b)$, $a < b, c < d \Rightarrow a + c < b + d$ and if one assume that $\beta^2 = 2s_1^2 b^2$ for some $b \in \mathbb{R}$, the following inequality is obtained;

$$\frac{1 - 4\gamma + 3\gamma^2}{\beta^2} > \frac{\gamma}{2s^2}$$
$$\Downarrow$$
$$1 - 4\gamma + 3\gamma^2 > b^2\gamma.$$

Unfortunately this equality can not be solved for all $\gamma \in]0, 1/2[$. The left side is less than zero for $\gamma \in [1/3, 1/2[$, while the right side is strictly positive for all $\gamma \in]0, 1/2[$. This implies that either there exists no function of the form $D(t, \beta) = e^{-t^2/\beta^2}$ which establishes the correct connections by Eq.(5.1.1), or that some of the above inequalities are too strict. These questions will not be further investigated. Instead the unsolvability of the above inequality will be discussed for some values of b. If $b = 1$ the above inequality holds for all $\gamma \in]0, 0.232[$, if $b = 1/2$ the above inequality holds for $\gamma \in]0, 0.3[$ and if $b = 1/4$ the above inequality holds for all $\gamma \in]0, 0.3233[$. By St.(5.1.1) this implies that if A is larger than resp. 1.357, 1.115 and 1.07 the correct connections are made by Eq.(5.1.1). Equivalently, if $d(x_1, x_0)$ is less than resp. 0.232, 0.3 and 0.323 the correct connections are made regardless of $A > 1$. It may be tempting to choose b to be very small in order to ensure that the correct connections are made for as low A as possible. For small values of b the effective width/support of $D(t, \beta)$ becomes narrow. This may cause problems in numeric implementations. b should be chosen such that $D(t, \beta)$ is "numerically" strictly larger than 0 for all $t \in]0, 1[$. In addition it is very little to gain by choosing $b < 1/4$. For the representation illustrated in Fig.(5.3), the value of b is 1. It appears that the correct connections are made.

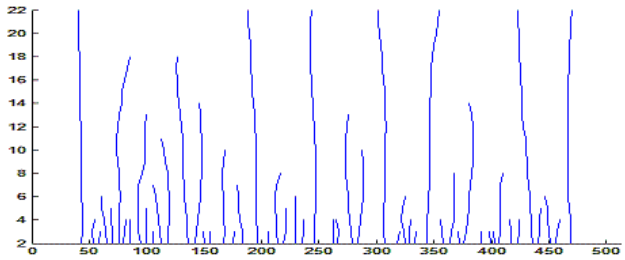
A summary of the proposed algorithm can be formulated as;

1. Compute $Wf(u, s)$ for $s \in \{s_1, ..., s_N\}$ and find the modulus-maximum.
2. Use Prop.(5.1.1) and Prop.(5.1.2) and find the possible connections of the modulus-maximum across scales.
3. If a modulus-maximum m at coarse scale can be connected with N modulus-maximum $\{m_i\}_{i=1}^N$ at the next finer scale, connect the modulus-maximum at coarse scale to the modulus-maximum at the finer scale which maximises $|Wf(m_i, s)D(d_i, s)|$. d_i is the spatial distance between m and m_i.

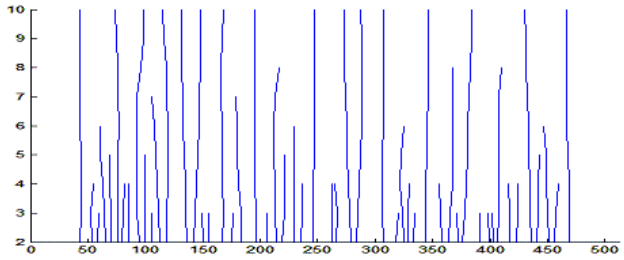
Note in step 3, that the modulus-maximum are traced from the coarse scales and "forced" to connect with something at the next finer scale. This guarantee that all modulus-maximum at coarse scales converge towards a modulus-maximum at the finest scale. This coincides with Th.(2.5.2). This would not necessarily be true if one used a fine-to-coarse tracking of the modulus-maxima. In addition the algorithm guarantees that one modulus-maximum corresponds to at most one modulus-maximum at the next finer scale. I.e. a "1-1" relation between modulus-maximum is guaranteed. *This will enable one to combine the good detection-performance at coarse scales together with the spatial accuracy at fine scales.* One can therefore construct edge-detectors with good detection- and localization-performance. Note that if a maxima-line should split, i.e. if two step-edges have similar strength, the proposed algorithm fails to include both modulus-maximum. The algorithm will force a 1-1 connection.

As discussed, the proposed algorithm fails to establish the correct connections if the strength of two step-edges is almost equal or equal. In addition, false connections may occur at small scales with the presence of noise. The relative part of the amplitude of the wavelet-transform corresponding to noise increase as the scale decreases. This may therefore cause false connections. On the other hand one may observe from Fig.(2.7) that the *sign* of modulus-maximum caused by noise is typically alternating. This will not prevent the possibility of false connections, but together with Prop.(5.1.2) this will restrict the spatial errors caused by eventual false connections. Another source of error is that one tries to make an algorithm designed for a model-function to hold for a general signal. For instance is there are several close edges in the signal. In the design of the algorithm, a pair of step-edges was used as the starting point. Because the Gaussian decays fast, this assumption will not cause errors if the edges are sufficiently isolated. If for instance one is dealing with three step-edges of similar sign with (almost) equal strength and they are equally close to each other, the behaviour of the maxima-line will be different than what discussed in Stat.(5.1.1). If the signal contains smoothened edges this may be another source of false connections. As discussed in Sect.(2.6.3), the wavelet-transform of a smoothened step-edge decays faster than a step-edge.

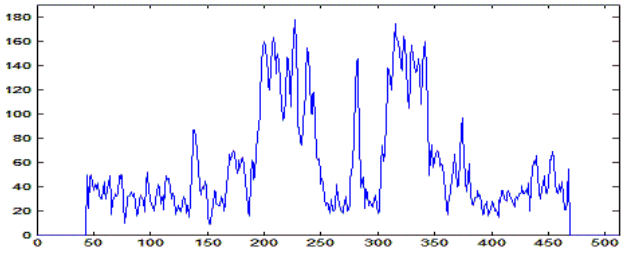
In order to verify the performance of the algorithm in practice, the algorithm is applied to the signal in Fig.(2.13). Fig.(5.3) illustrates the output of the algorithm. The input to the algorithm are all the modulus-maximum at scale $s = 2, 3, 4, ..., 22$. This is a much larger range of scales than what will be



(a)



(b)



(c)

Figure 5.3: The maxima-tree corresponding to the signal in Fig.(2.13). The proposed algorithm has been used to construct the maxima-tree. (a) $s \in [2, 10]$. (b) $s \in [2, 22]$. (c) The signal.

used in the discussion of the edge-detectors. This has been done to test the algorithm at both fine and coarse scales. In Fig.(5.3) there are no apparent false connections made by using this algorithm on the signal in Fig.(2.13).

After relating modulus-maximum across scales, one may assign to each modulus-maximum at the finest scale a quantity indicating at which scale it is first detected. *The length in time-scale plane of the maxima-line converging towards the modulus-maximum at the finest scale.* Assume that there are N

modulus-maximum m_i at the finest scale. Assign to each modulus-maximum m_i the class $[m_i]_{i=1}^N$ containing the modulus-maximum converging towards m_i. In this thesis, the length of the maxima-line which converges towards m_i is defined as the number of elements in $[m_i]$. In addition to the three remarks in the beginning of this section, one may now assign the following quantities to each modulus-maximum at the finest scale;

- The length $l_i = |[m_i]|$ of the maxima-line which converges towards m_i. $|\cdot|$ denote the number of elements.
- The weighted average amplitude $\bar{a}_i$ of the modulus-maximum which converge towards m_i for $i \in \{1, ..., N\}$.

For a modulus-maximum m at scale s, the weighted value of the wavelet-transform at the modulus-maximum is given by $s^{-1/2}Wf(m, s)$.

As mentioned in the introduction, the maxima-tree may be used as an (simple) edge-detector. The key-component of this edge-detector is the length of the maxima-lines. Note in Fig.(5.3a) that all but one of the seven modulus-maximum at the coarsest scale converges towards the most important features in the signal. The exception is the modulus-maximum which converges towards 440. Compared to the other edges, this edge is of little importance. This exception represents a unique characteristic feature of this edge-detector, namely its (relative) invariance to the of the strength edges. A small edge will have a long maxima-line if there are no stronger edges in the proximity of the edge. This agrees with the observations made at the beginning of the section.

The second task is to compare the proposed algorithm to with a few alternative methods from the literature. The first algorithm is by Lu et al. and is found in [7]. This method is the motivation behind the proposed algorithm. Although they may appear similar, there are some major differences between the methods. One obvious difference is the function D. The function D is by Lu et al. defined as the reproducing kernel between the wavelets at the two scales, say s_1 and s_0. If the wavelet is given by ψ_s^1, the corresponding expression of $D(t, s_0, s_1)$ is given by;

$$D(t, s_1, s_0) = Ke^{-1/2\frac{\tau^2}{{s_1}^2+{s_0}^2}} \left(\tau^2 - {s_1}^2 - {s_0}^2\right),$$

where K is some constant. Note that $D(t, s_0, s_1)$ has the form of the Mexican-Hat, and that the Mexican-Hat is not monotone. Monotony was the basic

assumption about the function $D(t, \beta)$ in the proposed algorithm. The algorithm of Lu et al. can be formulated as;

1. Compute $Wf(u, s)$ for $s \in \{s_1, ..., s_N\}$ and find the modulus-maximum.

2. Connect a modulus-maximum with the modulus-maximum at the *next coarser* scale which maximizes $|a_j^1 D(d_j, s_0, s_1)|$. a_j^1 is the amplitude of the modulus-maximum at the coarse scale, and d_j is the spatial distance between the modulus-maximum.

3. If a modulus-maximum can be connected with several modulus-maximum between two scales, use Prop.(5.1.1) to establish which modulus-maximum should be connected. If there is still not a 1-1 correspondence between modulus-maximum, connect the modulus-maximum which is most similar with respect to the amplitude and the spatial position.

In difference to the proposed method, this method uses a fine-to-coarse tracking of the modulus-maximum, i.e. the method begins at the finest scale. This may fail to ensure that all modulus-maximum converges towards finer scale. This may be observed in Fig.(5.4). Therefore one may encounter that modulus-maximum at coarse scale does not correspond to any modulus-maximum at finest scale. This is not in accordance with Th.(2.5.2). Worse, the function $D(t, s_0, s_1)$ does not necessarily establish the correct connections in Case(3b). This phenomenon is causing that there are often several connections between maxima-line of similar sign in Fig.(5.4). This is caused by a too slow decay of the ratio $D(\epsilon, s_0, s_1)/D(1 - \epsilon', s_0, s_1)$ with respect to scale. This error may cause that the final maxima-line converging to a modulus-maximum, may be given shorter time-scale length than it actually has. This phenomenon may be observed in Fig.(5.4). The final difference is that this method can be more tricky to implement, in particular the second remark in step 3 of the algorithm may be tricky.

One commonly encountered method for establishing the maxima-tree, is by regarding the Lipschitz-regularity of the possible connections [4]. This method is based on the set of possible connections obtained after using Stat.(5.1.1) and Stat.(5.1.2). The first step is to estimate the Lipschitz-regularity between each possible connection. Then, if a modulus-maximum m can be connected with several modulus-maximum at the next finer scale, m should be connected with the modulus-maximum which gives a Lipschitz-value most similar to the Lipschitz-values of the maxima-line which converges

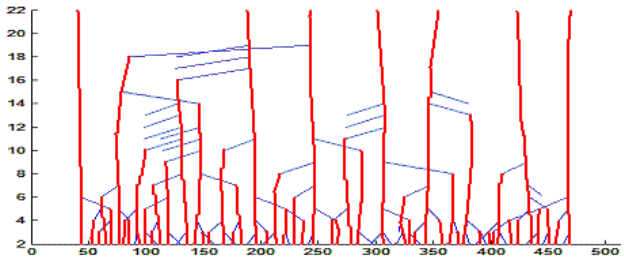


Figure 5.4: The red lines correspond to the maxima-tree corresponding to the signal in Fig.(2.13) by using the algorithm of Lu et al. for $s \in [2, 22]$

towards m. One may wonder how this is possible at the coarsest scales. In practice it is typically not a problem to relate modulus-maximum at the coarsest scales based on the *sign* and spatial position. It may however be tricky at fine scales. This method is efficient and reliable for manually establishing the correct connections, and may be more reliable for avoiding false connections at scales where the noise appears. Another advantage with this method is that is not coloured by assumptions made about the function or signal which is analysed, i.e. one does not need to use a function on the form $f(t) = u(t) + A(t)$ as a starting-point. The algorithm should in practice be reliable for all signals. On the other hand this method is trickier to implement. One may observe from Fig.(2.16) that a "similar" decay does not necessarily imply that the decay is (almost) constant. "Similar" decay should be understood as the decay which preserves a "similar behaviour" as the decay towards a modulus-maximum.

The latter and the proposed connection algorithm will return similar maxima-tree for the signal in Fig.(2.13) and other similar signals. The difference between these methods is the effort needed to implement the methods. The proposed method is by far simpler to implement. For the signal in Fig.(2.13) it is not necessary to use Prop.(5.1.2) to establish the correct connections. It suffices to find the modulus-maximum which maximises Eq.(5.1.1) of similar sign. This makes this method fast and very easy to implement. The Lipschitz-method is easy to use manually, and will be more reliable for a wider range of signals. The problem is that this method is trickier to implement.

5.2 2-D Maxima-Tree.

In this section different approaches for connecting modulus-maximum across scales in 2-D will be investigated.

In 1-D, the starting-point for the proposed connection-algorithm was some facts based on the *sign*. These facts hold for all signals. In order to obtain an 1-1 connection of modulus-maximum across scales, some assumptions about the signals were used to achieve the desired representation. In 2-D there are very few theoretic facts which can be used as a foundation of the algorithm. Instead assumptions are used to create the connection-algorithm.

The section will begin with discussing how one may connect single modulus-maximum across scales, before presenting a method for connecting line-segments across. This will enable to create 2-D edge-detectors which take advantage of information from several scales. One consequence of discussing line-segments is that one can obtain information about the size of the objects in the image. Why and how this can be done, is discussed in the last section.

5.2.1 Connecting modulus-maximum across scales in 2-D.

In this section a procedure which connects modulus-maximum across scales in 2-D will be presented.

A similar method as in 1-D will be used to relate modulus-maximum across scales in 2-D. Before this algorithm can be applied, one has to find the modulus-maximum which can be connected. In the previous section the *sign* of $Wf(u,s)$ was used to reduce the number of possible connections. In 2-D, the orientation of modulus-maximum will be used to reduce the number of possible connections.

The section begins with a brief summary of the information one may assign to each modulus-maximum, before discussing assumptions used to relate modulus-maximum across scales. At the end a connection algorithm will be presented.

The following information of the modulus-maximum in 2-D is available from the 2-D continuous wavelet-transform;

- The position of the modulus-maximum $\{((x_i^j, y_i^j), s_j)\}$ for $i \in \{1, ..., N_j\}$ and $s_j \in \{s_1, ..., s_J\}$. N_j is the number of modulus-maximum at a given

scale s_j.

- The *orientation* of a modulus-maximum, given by $Af((x_i^j, y_i^j), s_j)$.
- The *amplitude* of a modulus-maximum given by $Mf((x_i^j, y_i^j), s_j)$.

The definition of $Af((x, y), s)$ and $Mf((x, y), s)$ is found in Sect.(3.2). In the previous section, Th.(2.5.2) and two consequences of Prop.(2.5.1) were used to find which modulus-maximum could be connected across scales. As discussed in Sect.(3.3), there are some uncertainties of the structure of the 2-D modulus-maximum plane. It is uncertain whether all modulus-maximum converge towards finer scales, and eventually how a modulus-maximum converges. In this thesis it will be assumed that all modulus-maximum converges towards the finest scale. The relation between modulus-maximum across scales is an important property when analysing noisy signals. As discussed in Sect.(4.2), it is difficult to achieve both good detection- and localization-performance without relating modulus-maximum across scales. In order to relate modulus-maximum across scale, some assumptions are used of how modulus-maximum can be related across scales. The pay-off for relying on assumptions only, is that this connection-procedure is more vulnerable for false connections.

First, one needs to find the possible connections of modulus-maximum in 2-D. Assume that $\{((x_i^N, y_i^N), s_N)\}$ is a modulus-maximum as scale s_N. One has to decide which modulus-maximum at the next finer scale this modulus-maximum may be connected with. A modulus-maximum in 2-D is typically one of several modulus-maximum which constitute a connected line-segment. The first obstacle is to decide which of the modulus-maximum in the line-segment at the next finer scale, the modulus-maximum can be connected with. Fig.(5.5) illustrate the problem of determining which modulus-maximum should be connected across scales. It will not be discussed which of the possible connections that are more correct. In this thesis the possible connections at the next finer scale will be restricted to the modulus-maximum which intersects the line given by $\mathbf{m} + \mathbf{n}t$. $\mathbf{m}$ is the position of the modulus-maximum at coarse scale, and $\mathbf{n}$ is the direction of the modulus-maximum. In Fig.(5.5) this is illustrated by a solid red line. Note that the assumption is not reflecting a theoretic fact. This assumption will not guarantee that the correct modulus-maximum are connected across scales, nor will it guarantee that all modulus-maximum converge towards finer scale. The problem with convergence is caused by the possibility that the assumption does not

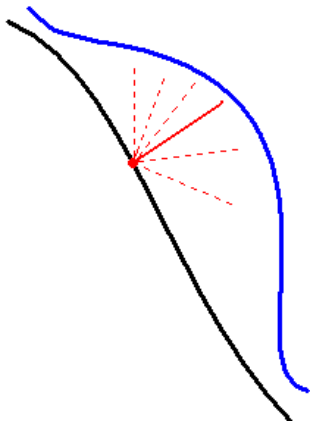
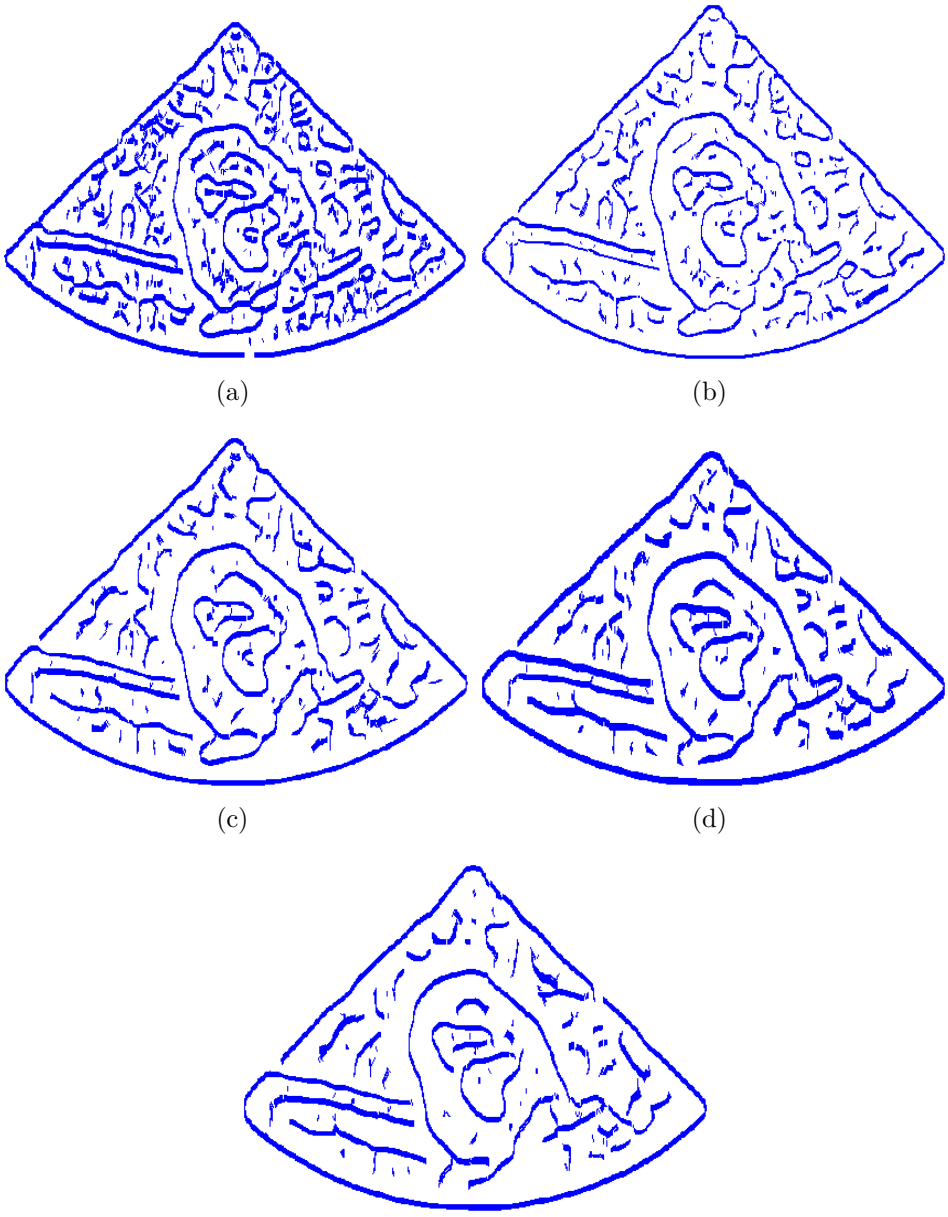


Figure 5.5: The black line corresponds to scale s_N and the blue line corresponds to scale s_{N-1} with $s_N > s_{N-1}$. The red dot is a modulus-maximum at coarse scale. The red lines show the possible connections to modulus-maximum at fine scale.

find any candidate modulus-maximum at the next finer scale. One should in particular be careful for false connection at the end of line-segments, or if a line-segment breaks up between two scales. Assume the modulus-maximum connected at fine scale are close to the modulus-maxima which should have been connected and a member of the same line-segment. Continuity of the wavelet-transform ensures that such false connections will not seriously affect the estimate of the Lipschitz-regularity. If the modulus-maxima at the next finer scale is on another line-segment, then a false connection will be critical. However, if the critical false connections are few and appear spuriously, these error-connections will not seriously corrupt the "visual" output of the edge-detector.

In 1-D, the *sign* of the modulus-maximum was used to reduce the number of possible connections between two scales. The corresponding quantity in 2-D is the *orientation* of a modulus-maximum. Unfortunately one does not have the possibility of establishing equally strict statements about the orientation of a modulus-maximum as the *sign* did in 1-D. However, if the signal is sufficiently nice and the CWT is computed for a sufficiently dense set of scales, one may say that the orientation of a modulus-maximum should not "change too much". The procedure of finding the candidate connections across scales can be formulated as follows;



(a) (b) (c) (d) (e)

Figure 5.6: The 2-D maxima-tree for the US-image in Fig.(1.4).

- Given two scales s_0 and s_1 with $s_0 < s_1$, and a modulus-maximum $\mathbf{m}$ with orientation $\mathbf{n}$ at scale s_1. The candidate connections at scale s_0 are the modulus-maximum which intersects the line given by $\mathbf{m}+\mathbf{n}t$ for $t \in [-\Omega, \Omega]$, $\Omega \in \mathbb{R}_+$. The orientation of $\{\mathbf{m}_i\}_{i=1^N}$ should be similar to the orientation of $\mathbf{m}$. N is the number of possible connections at the next finer scale.

The orientation-condition should be chosen loose enough to allow that the orientation changes as the details of the contour increases. One obstacle in 2-D, is that if an edge of small strength is (almost) perpendicular to an edge of large strength, the modulus-maxima corresponding to the two edges will be close at all scales. In 1-D an edge of large strength would prevent the existence of modulus-maxima corresponding to small edges at coarse scales. Therefore the orientation-condition should be chosen sufficiently strict to prevent false connections caused by edges with a different orientation. $\Omega > 0$ should be chosen large enough to allow that the spatial accuracy of the contour increases. On the other hand, Ω should be chosen small enough to reduce the likeliness of false connections, e.g. at the end of line-segments or in case of breaking of the line-segments. Without the Ω-condition, the algorithm would "force" a connection regardless of the spatial position of the modulus-maxima at fine scale. The connections made by this statement are typically redundant. Similarly as in 1-D, this redundancy will be achieved removed by the next step of the algorithm.

In the previous section a connection algorithm based on the spatial position of the modulus-maximum and the amplitude was investigated. It will be attempted to reduce the 2-D problem to a 1-D problem, and then use the algorithm from the previous section. In 2-D the discussion will be restricted to its simplest case, two parallel step-edges. Assume that the 2-D signal is given by $f(x,y) = u(x) + Au(x-1)$ for some $A > 1$. With respect to the semi-wavelet $\psi = \psi^1(x)\theta(y)$ and $\mathbf{v} = (v_1, v_2)$, the expression for the continuous wavelet-transform becomes;

$$\begin{aligned} W^{\mathbf{n}} f(\mathbf{v}, s) &= W^x f(\mathbf{v}, s) \\ &= \int_{-\infty}^{\infty} \theta_s(y)\, dy \Big(\int_0^{\infty} \psi_s^1(x - v_1)\, dx + A \int_1^{\infty} \psi_s^1(x - v_1)\, dx \Big). \end{aligned}$$

I.e. one may use a similar algorithm as in the previous section with $Wf(\tau, s)$ replaced with $Mf((x,y), s)$ and $d((x_1, x_2), (y_1, y_2)) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.

The procedure for connecting modulus-maximum across scales can be summarized by the following algorithm;

1. Compute $W^x f((x,y),s)$, $W^y f((x,y),s)$, $Mf((x,y),s)$ and $Af((x,y),s)$ for $s \in \{s_1, ..., s_N\}$.

2. Find local maxima of $Mf((x,y),s)$ in the local direction given by $Af((x,y),s)$ for $s \in \{s_1, ..., s_N\}$.

3. Find the set of possible connections between modulus-maximum across scales based on spatial-position and the orientation of the modulus-maximum.

4. If a modulus-maximum $\mathbf{m}$ at coarse scale can be connected to N modulus-maximum $\mathbf{m}_i$ at the next finer scale s, connect $\mathbf{m}$ to the $\mathbf{m}_i$ which maximizes $|Mf(\mathbf{m}_i, s)D(d_i, s)|$. d_i is the spatial distance between $\mathbf{m}$ and $\mathbf{m}_i$.

After this algorithm has been applied, one may for each modulus-maximum $\mathbf{m}_i$ at the finest scale define the class $[\mathbf{m}_i]$. The class $[\mathbf{m}_i]$ contains the modulus-maximum which converge towards $\mathbf{m}_i$. Assume there are N modulus-maximum $\mathbf{m}_i$ at the finest scale. The following quantities can now be assigned to each modulus-maximum at the finer scale;

- The length $l_i = |[\mathbf{m}_i]|$ in time-scale plane of the maxima-line converging to $\mathbf{m}_i$ for $i \in \{1, ..., N\}$.

- The weighted average amplitude $\bar{a}_i$ of the modulus-maximum along the maxima-line converging to $\mathbf{m}_i$ for $i \in \{1, ..., N\}$.

The uncertainty related to the behaviour of the modulus-maximum surface in 2-D, causes that this connection algorithm should always be followed with a visual investigation to ensure that there are no spurious connections. In practice, it suffices that the majority of connections are correct. A few isolated spurious connections will not seriously corrupt the visual output.

5.2.2 Connecting line-segments across scales in 2-D.

In this section convergence of line-segments across scales will be discussed.

In the previous section, the uncertainty of how a single modulus-maximum converges in 2-D was discussed. In order to circumvent the problem of how

single modulus-maximum converges in 2-D, one can consider convergence of line-segments. This will avoid some of the problems which were discussed in the previous section. Unfortunatly, this method will introduce new obstacles. Some of these obstacles, and how they are tried avoided in this thesis will be discussed. A consequence of dealing with line-segments rather than single points, is that one can extract information of the length (within a scale) of line-segments. This quantity may (optimally) be used to represent the circumference of an object. This will prove a very useful tool for identifying the important objects in the US-signal in Fig.(1.4). In Ch.(7) edge-detectors taking advantage of the convergence of line-segments will be discussed. It will be illustrated that these edge-detectors will be particularly useful for analysing the US-image. Although the connection-algorithm can be unstable and ad hoc, it is able to visualise the effect of taking advantage of the convergence of line-segments across scales.

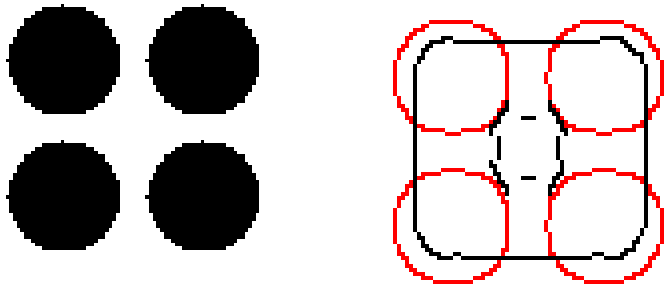
The section begins with an illustration of the problem of disconnected line-segments and some of the problems with convergence of line-segments. The underlying idea of how disconnected line-segments have been connected across scales in this thesis will be presented.

Fig.(5.8b) and Fig.(5.7) illustrate resp. the modulus-maximum at two scales from the image in Fig.(5.8a) and a segment of the signal in Fig.(1.4). Three phenomenons may be observed. First, from Fig.(5.7) one may observe that modulus-maximum are typically chained together into connected line-segments, representing the boarder of the objects in the signal. The number of isolated modulus-maximum is relatively small. Second, the length of the line-segment typically reflects the "size" or circumference of the object. The modulus-maximum of major objects are typically chained together into long line-segments, while modulus-maximum corresponding to noise, texture and smaller objects are chained together into shorter line-segments. The third observation is that connected line-segments have a tendency to break up into smaller line-segments as the scale decreases. The first and second observation represents how one may use this discussion to separate noise from the edges in the signal. If one could exploit these observations one would possess a powerful tool for identifying major objects in the image. The third observation represents the obstacle.

One of the goals of this section is to connect line-segments which encircle (a part of) an object. Recall the behaviour of SNR across scales; at coarse scale it was easy to find the important edges, while at fine scale the spatial

(a) (b)

Figure 5.7: (a) The edges of the signal in Fig.(1.4) at a fine scale. (b) The edges of the signal in Fig.(1.4) at a coarse scale. One may observe that some of the connected maxima-lines appear to represent the same object in Fig.(1.4), but with a different resolution.



(a) (b)

Figure 5.8: Illustration of an obstacle in relating line-segments inside a scale in 2-D. (a) The signal. (b) The red lines correspond to fine scale, and the black lines correspond to coarse scale.

accuracy was good. This motivated the construction of the maxima-tree. From Fig.(3.3) it may appear that a similar phenomenon exists regarding

the length of line-segments. One can easily find the important edges based on the length of the line-segments at coarse scales, but the spatial accuracy is poor. At fine scale the spatial accuracy is good, but because of the splitting of line-segments, it is more difficult to find the important objects based on the length of the line-segments.

Fig.(5.7) illustrates how a line-segment encircling an object breaks up into several line-segments between two scales. In Fig.(5.8) it is illustrated how one line-segment break up to encircle four objects. Recall from the previous sections that in the construction of 1-D and 2-D maxima-tree one tried to achieve a 1-1 correspondence between modulus-maximum across scales. This was desirable in order to construct an edge-detector with both good detection- and localization-performance. In the case of convergence of line-segments, a 1-1 correspondence is desirable in order to benefit from length of line-segments at coarse scales. The problem is that this is not necessarily justifiable. Assume one line-segment encircling an object breaks into several line-segments encircling the same object as in Fig.(5.7). In this case a 1-1 correspondence can be justified. A 1-1 correspondence can be achieved by connecting all the line-segments at fine scale which originate from one line-segment at coarse scale. However, if one line-segment at coarse scale breaks up into several line-segments encircling several objects, one may encounter problems with justifying a 1-1 correspondence. For instance in the signal in Fig.(5.8) one would like to be able to separate the four circles as different objects. Since they all originate from the same line-segment they would in case of a 1-1 correspondence not be separable. This problem of convergence in will not be addressed in this thesis. In order to reduce (but not eliminate) this problem, one may select the range of scales with care. For instance for the US-image in Fig.(1.4) one may see in Fig.(3.3) that the interesting objects remains in shape for $s \geq 4$. For small scales it loses some of its structure.

It this thesis, connecting line-segments across scales is based on the following statement in [7];
Two line-segments should be connected if they are separated by a small gap, and the corresponding line-segment at the next coarser scale is connected.
Two line-segments which have been connected, is said to be in the same class of line-segments which have been connected inside a scale. The first criteria; that the gap between two line-segments should be small, will help to avoid some false connections caused by the phenomenon illustrated by Fig.(5.8). How small the gap should be allowed is uncertain. In this thesis an appropriate upper limit for the gap has been found by experimenting. The cor-

responding connection-procedure using the estimated limit for the gap, has been successful for the US-image in Fig.(1.4) and the US-images displayed in App.(A). The last question is how to find the "corresponding line-segment" at the next finer scale. If the connection-procedure for modulus-maximum discussed in the previous section were fail-safe, this would be a relatively easy task. One could connect all line-segments at the next finer scale, towards some modulus-maximum of the line-segment at coarse scale converges. Unfortunately such an algorithm will be vulnerable for false connections. Instead line-segments are connected across scale based on the following observation in Fig.(3.3). *Two corresponding line-segments across scales have some points in common, i.e. some modulus-maxima at both scales have equal spatial position.* With this observation, a line-segment at fine scale is related to a line-segment at the next coarser scale, if they have some modulus-maxima which have the same spatial position. Although, this connection-procedure may appear unreliable, it has been "fairly" successful with respect to analysis of the US-image in Fig.(1.4) and similar medical US-images displayed in App.(A). Note that this observation is supported by the fact that at large scales the contours in the image behaves as step-edges. The average spatial position of a line-segment at two scales will therefore be similar. This increase the likeliness of intersection of two corresponding line-segments across scales.

This procedure is unstable for $s \leq 3$. At fine scales the density of the modulus-maximum will cause false connections. In addition, all edges at fine scale begin to behave as Dirac-edges. The (average) spatial position of a Dirac-edge will change across scales. This is not a problem with respect to the main-purpose of this thesis; to achieve a representation which can be used for medical applications. For this purpose it is not necessary to use the modulus-maximum at scales less than $s = 5$. In one case on is certain that this algorithm will break down, if there is an edge on the form $f(x, y) = \delta(x)$ in the image. In this case there will for each y be two modulus-maximum, see Ex.(3.1.2). The spatial-position of these modulus-maximum will not be equal at different scales, i.e. they will not have any modulus-maximum in common at two different scales. This is not a problem for the US-image in Fig.(1.4), but this caused problem when attempting to analyse a CT-image, see Fig(A.5). In this case the "skin" behaves as the function $f(x, y) = \delta(x)$.

One should use this algorithm both from coarse to fine and from fine to coarse scales. It may occur that a line-segment are disconnected at coarse scale and connected at fine scales.

After connecting line-segments inside and across scales, some additional quantities about the line-segments can be introduced;

- The length $\{L_k^s\}$ (inside a scale) of connected (inside scales) line-segments, $k \in \{1, ...N_s\}$ and $s \in \{s_1, ..., s_N\}$. N_s is the number of distinct classes of connected line-segments at scale s.

- The average amplitude $\bar{A}_k^s$ of the connected line-segments, $k \in \{1, ..., M\}$ and $s \in \{s_1, ..., s_N\}$. N_s is the number of connected line-segments at scale s.

- The length $\{l_k\}$ $k = 1, ..., N$ (across scales) in the time-scale plane of the line-segments converging towards a (connected) line-segment at the finest scale. The length is the number of scales there exist a line-segment converging to the line-segment at the finest scale. N is the number of (connected) line-segments at the finest scale.

- The weighted average amplitude $\bar{a}_k^s$ $k = 1, ..., N$ of the line-segments converging to a line-segment at the finest scale. N is the number of connected line-segments at the finest scale.

The 1.st quantity does not necessarily correspond to the circumference of an object. If the boarder of an object at coarsest scale is constituted of several line-segments, the set of related line-segments at the finest scale do not correspond to the circumference of the object. It is however not necessary that the related line-segments correspond to the circumference, but the longer the line-segments the easier to distinguish noise from edges. If eventually the circumference is obtained this is a nice bonus.

The uncertainty of how line-segments (and points) converge in 2-D is unfortunate. One is not guaranteed that correct connections are made. If however the correct connections of line-segments are done across and inside scales, one has a powerful tool for identifying the important edges in the signal. This will be illustrated in Ch.(7), where edge-detectors taking advantage of these properties will be introduced. Without destroying the excitement the writer can reveal that edge-detectors taking advantage the behaviour of either line-segments or modulus-maximum across scales, are superior to edge-detectors which depend on the modulus-maximum at a single scale only. This holds in particularly for noisy signals, such as US-image in Fig.(1.4). In this thesis, the output of the edge-detectors will always be visually investigated.

One will be able to identify false connections if they appear in the output. If they do not appear in the output, there is no need for further investigation, at least for the purpose of this thesis.

Chapter 6

1-D Edge-Detectors.

In this chapter the 1-D edge-detectors will be presented.

This far in the thesis, properties of the CWT and the modulus-maximum have been discussed from a relatively theoretic viewpoint. It is time to investigate how these properties can be combined in order to locate the important edges in a 1-D signal. The signal which will be investigated is the 1-D signal in Fig.(2.13). This signal is a 1-D ray of the US-image in Fig.(1.4).

Three 1-D edge-detectors will be discussed; one single-scale and two multi-scale edge-detectors. This will be used to visualise differences between single- and multi-scale edge-detectors with respect to noisy signals. The motivation behind the edge-detectors are found throughout the thesis; Stat.(2.1.1) and the section about SNR is the basis for the single-scale detector. The maxima-tree, and the discussion about the Lipschitz-regularity are the foundations for the multi-scale edge-detectors.

6.1 Probability & Amplitude Thresholding.

The first edge-detector is the simplest and the most obvious; to threshold the amplitude of the modulus-maximum. This can be done by estimating the probability-distribution of the noise in the signal, e.g. by using a Wiener-filter.

In Stat.(2.1.1) it was verified that the absolute-value of the wavelet-transform is large in neighbourhoods where the signal changes much and with a long duration. In this section an edge-detector taking advantage of this will be presented. The underlying idea of why this will separate the noise from

the important edges, is that the important edges typically changes much and with a long duration. On the other hand, noise changes little and with a short duration. One advantage with the presented method is that a threshold-value can be obtained automatically. In this case, the edge-detector requires no collaboration with the operator. Another advantage of this method is that in its simplest form only require the computation of the wavelet-transform at a single scale. This reduces the amount of computations, and therefore the method is fast. One can also use this method without estimating the noise; by guessing a threshold. In this case the method is no longer automatic.

In Sect.(4.2), SNR was introduced as a quantification of the likeliness of a modulus-maximum to correspond to an important edge. The larger value of SNR at a modulus-maximum, the more likely that it corresponds to an edge. In this section it will be discussed how big SNR of a modulus-maximum has to be in order to represent an important edge. It will be discussed and illustrated that the method using a noise-estimate is not adequate with respect to the type of signals which is analysed in this thesis. A lot of redundant information is present in the output. Representations closer to the optimal are found by guessing a value for the threshold.

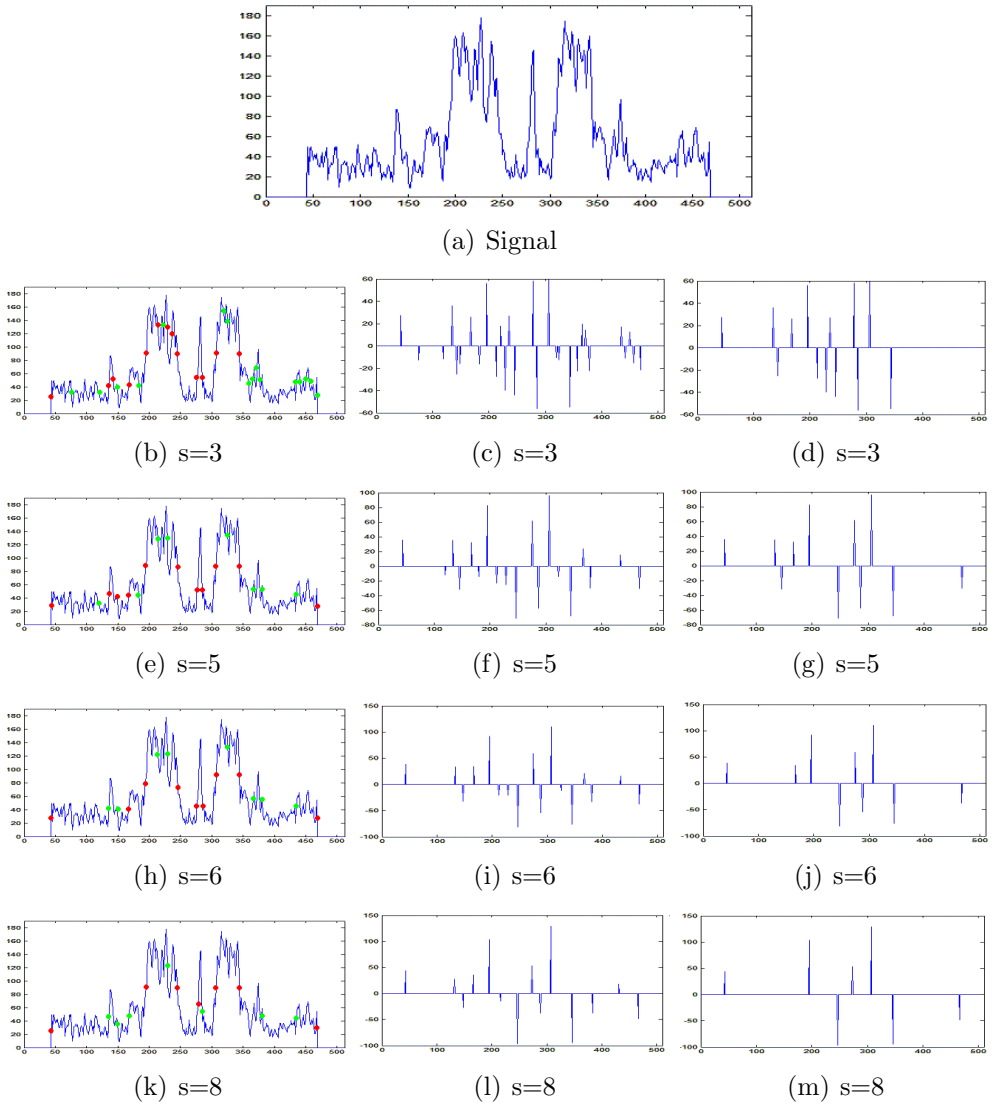
The section begins with a discussion of how the threshold can be estimated, before presenting a single- and multi-scale algorithm for detecting the important edges in a signal. At the end, the output of the edge-detector will be discussed.

A modulus-maximum will be rejected to represent an important edge based on its probability to correspond to noise. Assume the variance of the noise is σ^2, and that the noise is Gaussian distributed. The probability-distribution of the noise is given by;

$$p(a) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{a^2}{2\sigma^2}}.$$

The aim is to reduce the probability of including a false edge. Assume that candidate edges with probability less than some probability threshold $P_{\%}$ are rejected to represent an important edge. This implies that all modulus-maximum with amplitude less than $A_{\%}$ is assumed to be noise, where $A_{\%}$ is the solution of;

$$\int_{-A_{\%}}^{A_{\%}} p(t)\, dt = P_{\%}.$$



(a) Signal

(b) s=3 (c) s=3 (d) s=3

(e) s=5 (f) s=5 (g) s=5

(h) s=6 (i) s=6 (j) s=6

(k) s=8 (l) s=8 (m) s=8

Figure 6.1: The edges detected by using a threshold for the amplitude of the modulus-maximum. (b)-(e)-(h)-(k): The red dots indicate edges in common for both representations. The green dots indicate additional edges in the representation obtained by using noise-estimate. In (c)-(f)-(i)-(l) the modulus-maximum kept by using the threshold obtained by using noise-estimate. In (d)-(g)-(j)-(m) the modulus-maximum kept by using a guessed value of $A_{\%}$.

In Sect.(4.2), SNR was introduced as a quantity to measure the likeliness of a modulus-maximum to correspond to an important edge. The larger value of SNR at a modulus-maximum, the more likely that it corresponds to an important edge. It was not discussed how large the value of SNR had to be at a modulus-maximum in order to be considered as an important edge. From above, the strength of an edge must be above $A_{\%}$ to avoid being refused as noise. Combined with the expression of SNR, the threshold for the modulus-maximum becomes;

$$T_{\%} = SNR(A_{\%}) = \frac{A_{\%}}{\sigma} \frac{\int_0^{\infty} \psi_s(t)\, dt}{\| \psi \|_2} = \frac{A_{\%}}{\sigma} \Sigma(\psi_s) = \frac{A_{\%}}{\sigma} \sqrt{s} \Sigma(\psi).$$

If a modulus-maximum m should be classified as an important edge, then;

$$\frac{|Wf(m,s)|}{\sigma \| \psi \|_2} > \frac{A_{\%}}{\sigma} \sqrt{s} \Sigma(\psi) = T_{\%}.$$

The amplitude of a modulus-maximum corresponding to a smooth step-edge decays faster than for a step-edge as the scale decreases. A smooth step-edge represents a slow change of state in the signal. By Th.(2.6.1) it follows that the wavelet-transform with respect to the wavelet ψ^1 in Eq.(2.1.4) of a smooth step-edge, decays faster than for a step-edge. This may cause some edges to be rejected at fine scales. In Sect.(5.1), a method for relating modulus-maximum across scales was discussed. The result was for each modulus-maximum m_i, the class $[m_i]$ of the modulus-maximum which converge towards m_i. A suggested multi-scale approach to detect the edges in a signal, is to include all modulus-maximum m_i at the finest scale which satisfy that some elements of $[m_i]$ is above the threshold.

The edge-detector can be summarized by the following algorithm:

1. Compute $Wf(u,s)$ for $s \in \{s_1, ..., s_J\}$.

2. Locate the modulus-maximum m_i^s at each scale. If $J > 1$, connect modulus-maximum across scales. Let $[m_i]$ for $i \in \{1, ...N\}$ denote the set of modulus-maximum which converges towards m_i. N is the number of modulus-maximum at finest scale.

3. Estimate the noise in the signal, e.g. by using Wiener-filtering presented in Sect.(4.3).

4. Decide a threshold of the probability for an edge to correspond to noise, typically $P_{\%} \in [80, 100]$. Find the lower bound $A_{\%}$.

5. Calculate $T_{\%,s} = \frac{A_{\%}}{n_0}\Sigma(\psi_s)$.

6. - **(Single scale):** If $SNR(m_i^1, s_1) > T_{\%,s_1}$, the modulus-maximum m_i is included as an edge in the signal, otherwise rejected.
   - **(Multi scale):** If $SNR(m_i^j, s_j) > T_{\%,s_j}$ for any $m_i^j \in [m_i]$ the modulus-maximum m_i is included as an edge in the signal, otherwise rejected.

Without using a noise-estimate one should skip Step 3 and guess a value of $A_{\%}$ in Step 4.

Consider Fig.(6.1) demonstrating the algorithm applied to the signal in Fig.(6.1,a) with and without a noise-estimate. First one may observe that the algorithm using a noise-estimate is not appropriate for the purpose of this thesis. For the analysis in this thesis, one is not only interested in removing noise, but also modulus-maximum corresponding to texture and less important edges. Without a noise-estimate, one may observe that at fine scales there are still unwanted edges in the representation. These are edges of short duration, i.e. they are felt as Dirac-edges. The Lipschitz-regularity is approximately -1 for such edges. This implies that the amplitude increases as the scale decreases, in difference to step-edges for which the amplitude decreases across scales. Observe for instance in Fig.(6.1c) that one can not remove the modulus-maximum at approx. 150 without removing the modulus-maximum at 50. The first one could do without, but the second is an edge one would like to preserve. Similarly, one may observe that such a separation is no problem at coarse scales, but at coarse scales it is more difficult to detect the edge at 270. This coincides with the discussion of the uncertainty-principle of edge-detection in Sect.(4.2). Another observation is that there is little to gain by using the multi-scale approach. In the representation without using a noise-estimate, the edge at $u = 460$ is not included at small scales. In Sect.(2.7) it was calculated that this edge was Lipschitz-0.41, i.e. it behaves as a smoothened edge. This causes the amplitude of this edge to decrease faster and is therefore discriminated at fine scales. The disadvantage of this method of using multi-scale information, is that it only includes and not excludes modulus-maximum. If one begins at a too fine scale, one is not be able to remove the "unwanted" edges which often appear at fine scales. In

addition it only exploits information from one single scale to decide which edges should be included. It does not take into account the behaviour of a modulus-maximum across scales. This is why this edge-detector is treated as a single-scale edge-detector.

All but the first of the representations in Fig.(6.1,d,g,j,m) gives an acceptable representation of the signal. With respect to the preliminary targets defined in Sect.(1.2.2), the representation in (g) is closest to the optimal. One may note that for finding the optimal solution, this edge-detector depends on two parameters; which scale to be used and the appropriate threshold for the amplitude. This makes it difficult to find a representation which is close to optimal.

The conclusion with respect to analysis of noisy signals, is that using a threshold of the amplitude of the modulus-maximum at a single scale is a pay-off between accuracy and the quality of the representation. At fine scales on is able to detect small objects in the signal, but there is a large presence of unwanted information. In addition an edge representing a smooth transition between tissues may not be detected at fine scales. At coarse scales one will detect the large objects in the signal without including unwanted information, and one is able to detect smooth edges. The disadvantage is that at coarse scales one will not detect small objects in the signal.

6.2 Multi-scale Edge-Detectors.

In this section two multi-scale edge-detectors will be presented.

In the previous section it was concluded that using the information of the modulus-maximum at a single scale only was not optimal with respect to the application of this thesis. In this section it will be investigated how one may exploit the behaviour of the modulus-maximum at several scales to locate the important edges in a 1-D signal. By using information across scales, one will be able to avoid the uncertainty of which scale should be used as the finest. For instance it was discussed in the previous section that one could not use amplitude-threshold at scales finer than $s = 5$ without including unwanted information. In this section the finest scale is $s = 2$.

The discussion begins with the underlying idea of the edge-detectors, and a brief summary of the properties used by the edge-detectors. Some interesting observations of the output will be discussed, and based on these observations some unique characteristics of each edge-detector will be iden-

tified.

The edge-detector which uses the Lipschitz-regularity (denoted Lipschitz edge-detector) of a modulus-maximum will be good for finding the points where the signal changes with a long duration, but less good for finding the points where the signal changes most. Another disadvantage with the Lipschitz edge-detector is that it can be difficult to find a thresholds which give an optimal representation. The second multi-scale edge-detector will be capable to find the points where the signal changes the most, and works well for removing unwanted information. It will be concluded that this is the most suitable for the analysis in this thesis of the edge-detector discussed in this thesis. The disadvantage with the detector is that one can not estimate or guess a threshold which gives an acceptable representation. On the other hand it is the detector from which it is easiest to find the threshold which gives a representation close to optimal.

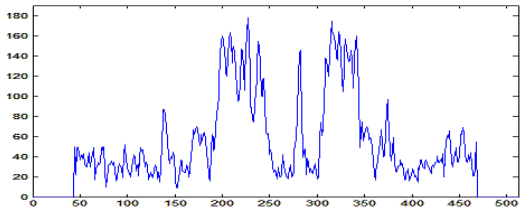
The discussion begins with a brief summary concerning the length of a maxima-line. The two edge-detectors and their characteristics will be discussed.

In Sect.(5.1) the length l_i of a maxima-line was defined as;

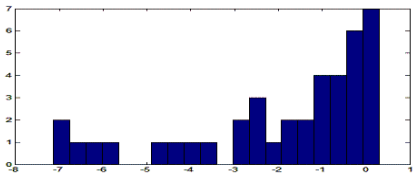
$$l_i = |[m_i]| = (\text{the number of modulus-maximum converging to } m_i),$$

for $i \in \{1, ..., N\}$, and N is the number of modulus-maximum at the finest scale. The maxima-tree in Fig.(5.3b) will be used for the discussion of the edge-detectors in this section. In this case the maxima-tree contains too few scales in order to be appropriate for edge-detection, see Sect.(5.1). Some additional information is needed in order to find the important edges in the signal.

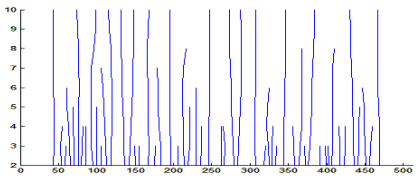
The underlying idea behind the Lipschitz edge-detector is that decay of modulus-maximum corresponding to noise differs from the decay for a step-edge. The difficulty is that the Lipschitz-regularity is invariant to the strength of the edge. By imposing a threshold on the length of the maxima-line, this edge-detector should be able to find the most important edges in a signal. The discussion in Sect.(2.6), and the example in Sect.(2.7) illustrates the necessity of a close investigation of each modulus-maximum in order to estimate which scales should be taken into account for estimating the local Lispchitz-regularity. A similar investigation is not realistic in general; it requires manual investigation, and will be time-consuming.



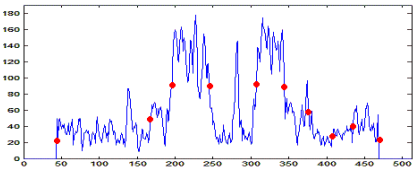
(a) Signal.



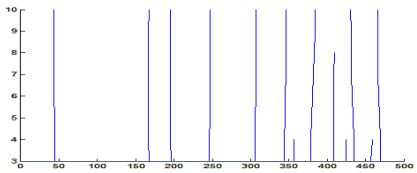
(b) Histogram of Lipschitz-values.



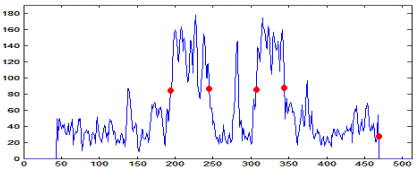
(c) Maxima-tree.



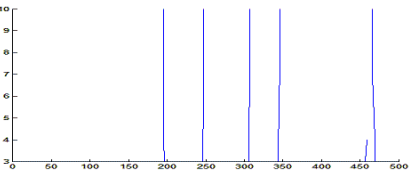
(d)



(e) The tree-branches with average Lipschitz-regularity larger than -0.6.



(f)



(g) The tree-branches with average Lipschitz-regularity larger than 0.

Figure 6.2: The edges detected by using the average Lipschitz-regularity of the maxima-lines. (b) The histogram of the average Lipschitz-regularity of the tree-branches. (e)-(g) Displays the tree-branches with average Lipschitz-regularity larger than resp. -0.6 and 0. (d)-(f) The points which are detected as edges. The modulus-maximum with the shortest maxima-line are not included to represent an edge.

The quantity used as a representation of the Lipschitz-regularity, is the average Lipschitz-regularity of the maxima-line which converges towards the modulus-maximum. The algorithm can be summarized as;

1. Compute the $Wf(u,s)$ for $s \in \{s_1, ..., s_N\}$. Find the modulus-maximum m_i^j for $i \in \{1, ..., N_j\}$ and $s_j \in \{s_1, ..., s_J\}$ where N_j is the number of modulus-maximum at scale s_j.

2. Find $[m_i]$ for $i \in \{1, ..., N_1\}$.

3. Compute the length $l_i = |[m_i]|$ of the maxima-line converging to the modulus-maximum m_i for $i \in \{1, ..., N\}$.

4. Compute the average Lipschitz-regularity of the maxima-line converging to m_i.

5. Remove the modulus-maximum with average Lipschitz-regularity smaller than T_{Lip} and length less than a given threshold. T_{Lip} is typically smaller than 0 and larger than -1.

An appropriate threshold for T_{Lip} can be difficult to determine. From the discussion in Sect.(2.6), the threshold will depend on influence of noise and smoothness of the edges in the signal. An appropriate threshold for the Lipschitz-regularity have in this thesis been found by experimenting with different values of $T_{Lip} \in]-1, 0[$. Experience has proved that $T_{Lip} \approx -1/2$ is a good starting-point for the signals analysed in this thesis.

Fig.(6.2e,g) illustrates the maxima-branches with average Lipschitz-regularity larger than -0.6 and 0. In Fig.(6.2d,f) the modulus-maximum with the longest maxima-lines have been marked as edges in the signal. Observe that the modulus-maximum which are kept with $T_{Lip} = 0$ and long maxima-line, are the 4 points where the signal changes most in addition to the edge at 469. As discussed in Sect.(2.7), the edge at 469 is felt as smoothened by the wavelet-transform. The 4 large edges are less influenced by close edges. In addition these edges are felt as smoothened by the wavelet-transform, such that the estimated Lipschitz-regularity increases. Observe that the step-edge at 50 is not included with $T_{Lip} = 0$. This is the edge one should expect to behave most as a step-edge. One can conclude that 0 is a too strict value of T_{Lip} for the signal analysed in this thesis.

One unique characteristic-feature of this edge-detector is its ability to remove edges of short duration. This is illustrated by the "instant" edge at

270. Such edges behave as Dirac-edges. Their corresponding Lipschitz-value is therefore approximately -1. Another characteristic-feature of this edge-detector is its "invariance" of the strength of the edge. This is illustrated by the small edges at 410 and 440 in Fig.(6.2d). As discussed in Sect.(5.1), the length of a maxima-line depends not only on the strength of the edge, but also on edges in its proximity. The Lipschitz-value of an edge is invariant of the strength. The consequence is that the corresponding edge-detector is relatively "invariant" to the strength of an edge.

Unfortunately none of the characteristic-features of this edge-detector are shown to an advantage with respect to the analysis of this thesis. The Dirac-edge at 297 represents one possible separation between the cysts in Fig.(1.4). This was not necessary to detect from a medical viewpoint, but it was one of the objects one would like to detect from a theoretic viewpoint. The invariance of the strength of an edge is not desirable in this thesis. One seeks the points where the signal changes the most. By exploiting the detection performance of the maxima-tree, see Sect.(5.1), one will be able to remove small edges by including coarser scales in the maxima-tree.

The amplitude of a modulus-maximum corresponding to a major edge is large at several scales, while a modulus-maximum corresponding to noise is large for a few scales only. The second multi-scale edge-detector takes advantage of this observation. In [7] this detector was introduced to be effective for finding the points of large change in a noisy signal. In Sect.(5.1), the average amplitude $\bar{a}_i$ of the maxima-line converging to m_i was defined as;

$$\bar{a}_i = (\text{weighted average amplitude of the modulus-maximum along the maximaline converging to } m_i).$$

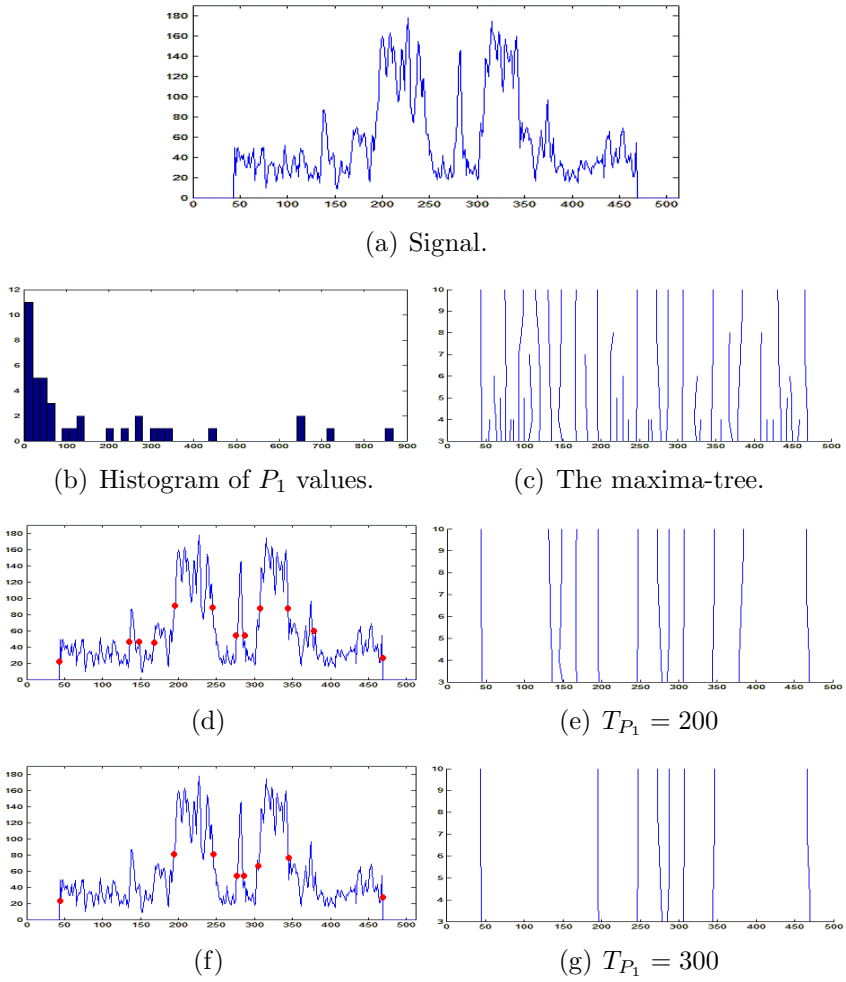
The algorithm assigns to each modulus-maximum at the finest scale the quantity,

$$P_1(m_i) = \bar{a}_i \cdot l_i.$$

Step 1 and Step 2 of this algorithm is similar with the latter algorithm. The algorithm can be summarized by the following steps;

3. Compute the average amplitude $\bar{a}_i$ of the tree-branch converging to m_i.
4. Compute $P_1(m_i) = l_i \cdot \bar{a}_i$ for $i \in \{1, ..., N\}$.

5. Find an appropriate threshold T_{P_1} separating important from less important modulus-maximum.



(a) Signal.

(b) Histogram of P_1 values.

(c) The maxima-tree.

(d)

(e) $T_{P_1} = 200$

(f)

(g) $T_{P_1} = 300$

Figure 6.3: The edges detected using P_1. (e)-(g) The tree-branches corresponding to modulus-maximum at $s = 2$ with P_1 values larger than resp. 200 and 300. (d)-(f) The points marked as edges.

Observe that this algorithm combines the length and the amplitude of a maxima-line into a single quantity to represent the modulus-maximum at the finest scale. A consequence is that a modulus-maximum with a long maxima-line is not necessarily included, or that a short maxima-line is excluded. P_1

arrange the modulus-maximum in a hierarchic structure. Modulus-maximum corresponding to a large edge with a long duration is assigned the largest values. The smallest values are modulus-maximum corresponding to edges of small strength which exists for a short duration, such as noise. A threshold is found by considering the values of P_1 in a histogram. By choosing an appropriate value for the threshold, denoted T_{P_1}, one should be able to obtain a nice representation of the most important edges in a signal.

Fig.(6.3) illustrates the algorithm applied to the signal in Fig.(6.3a). By comparing the difference between the two representations, one may observe the hierarchical structure of the values of P_1. Another observation is that the algorithm does not discriminate edges based on their Lipschitz-regularity; both Dirac-edges and step-edges are included in both representation. In difference to the Lipschitz edge-detector, this method will enable one to distinguish step-edges of small strength from step-edges with large strength. This method will detect smooth-edges, i.e. edges corresponding to a slow change of state in the signal. In the previous section it was observed that by using a threshold of the modulus-maximum, such edges would be discriminated at fine scales.

The disadvantage of this method is that there is no predefined value of the threshold. In the previously discussed methods, one could in advance estimate or guess some appropriate threshold which would give a useable representation of the signal. This is not as easy with this method. In Sect.(7.1.2) it will be discussed how one can find a natural threshold with this method. Because there are relatively few modulus-maximum in the 1-D case, this discussion is postponed. The values of P_1 corresponding to important edges are typically discretely distributed (depending on the interval used to make the histogram). This enables the operator to exclude edges until the optimal representation is obtained. For the histogram in Fig.(6.3g) one have in theory 8 possibilities of a threshold, assuming that all values below 100 is noise. In practice it requires at most 3 guesses to obtain a nice representation. For both the Lipschitz-method and in particular the method using an amplitude threshold one may in worst case use several attempts to find a good representation. In this detector it is only necessary to vary one threshold, while for two other detectors one may have to experiment with 2 different parameters in order to find the optimal representation. It is easier to find the optimal threshold for the latter edge-detector.

The representation obtained in Fig.(6.3e) contains all the edges of the signal in Fig.(6.3a). All the points which at the first glance appears to be an

edge in the signal, are represented as an edge. The representation obtained in Fig.(6.3g) is optimal with respect to the preliminary targets defined in Sect.(1.2.2), taken into account the position of the ray in the signal, see Fig.(2.13). Comparing with the preliminary target and the US-image one see that all the edges which one was interested in detecting, are marked as an edge.

With respect to the analysis of noisy-signals, the conclusion is that P_1 will detect edges of large strength corresponding to a change of state in the signal of both short and long durtation. In addition P_1 will detect an edge representing a slow change of state in the signal, i.e. a smooth edge.

The Lipschitz edge-detectors will detect edges corresponding to a change of state in the signal of long duration, "independent" of the strength of the edge. Edges corresponding to a change of state in the signal of short duration will not be detected, regardless of the strength of the edge.

Chapter 7

2-D Edge-Detectors.

This chapter contains a discussion about 2-D edge-detectors.

We have now reached the target of the thesis; to find the edges in the US-image in Fig.(1.4). The purpose of this section is to visualise how one can combine properties of the 2-D continuous wavelet-transform, and create an edge-detector which is suitable for the US-image.

The investigation follows a similar procedure as in the previous section. The discussion begins with investigating two single-scale edge-detectors, before introducing three multi-scale edge-detectors. Based on the output, some characteristics of each edge-detector will be identified. (Dis-)advantages of the edge-detectors with respect to the analysis of this thesis will be discussed. The preliminary targets discussed in Sect.(1.2.2) will be used to compare how close a representation is to being optimal.

Similarly as in the previous section, it will be concluded that the edge-detector which take into account the amplitude of modulus-maximum at several scales are more appropriate with respect to the analysis of this thesis. Another conclusion is that with respect to the application of this thesis, the edge-detectors which exploit information about line-segments are more suitable that the edge-detectors which only exploit the behaviour of a single modulus-maximum.

The chapter begins with discussing two single-scale edge-detectors which exploit the amplitude of a modulus-maximum, and the length and amplitude of a line-segment. Three edge-detectors which take advantage of the behaviour of modulus-maximum and line-segments across scales will be discussed.

7.1 Single-Scale Edge-Detectors.

7.1.1 Amplitude-Thresholding.

In this section it will be discussed how one can detect edges in an image by using a threshold of the amplitude of a modulus-maximum.

The algorithm which will be presented is a single-scale version of a more general multi-scale edge-detector developed by Canny in [3][1]. The advantage of the method is that it requires minimal computational effort, and is easy to implement. In addition it has the advantage that the value for the threshold can be pre-estimated.

The 2-D algorithm which will be presented is similar to the 1-D formulation in Sect.(6.1). The difference is that a two-threshold method is used, in [3] denoted *thresholding with hysteria*. Using a threshold with hysteria will help to preserve connected line-segments in the representation of the image.

First thresholding with hysteria is explained, before presenting an algorithm for detecting edges based on the amplitude of the modulus-maximum. At the end of the section, the performance of the edge-detector will be discussed with respect to the image in Fig.(1.4) and the preliminary targets defined in Sect.(1.2.2).

One problem which often occurs for edge-detectors based on a threshold of the amplitude of the modulus-maximum, is a phenomenon called *streaking*. The problem of streaking is discussed in [3]. Modulus-maximum of the CWT are often connected into a line-segment representing the boundary of an object. The problem which one may encounter is that the amplitude of the modulus-maximum varies along the line-segment. Streaking occurs if one part of the line-segment is above a threshold, while another part is below the threshold. This phenomenon may be observed in Fig.(7.1,b,e,h). This results in broken boundaries in the representation of the objects. One possible solution to avoid this problem is presented in [3], as *thresholding with hysteria*. Thresholding with hysteria uses two thresholds, one high and one low threshold. If a modulus-maximum is above the high threshold, it should immediately be marked as an edge. If a modulus-maximum is above the low threshold and connected to a line-segment which at some point is above the high threshold, it should also be marked as an edge. The algorithm can be

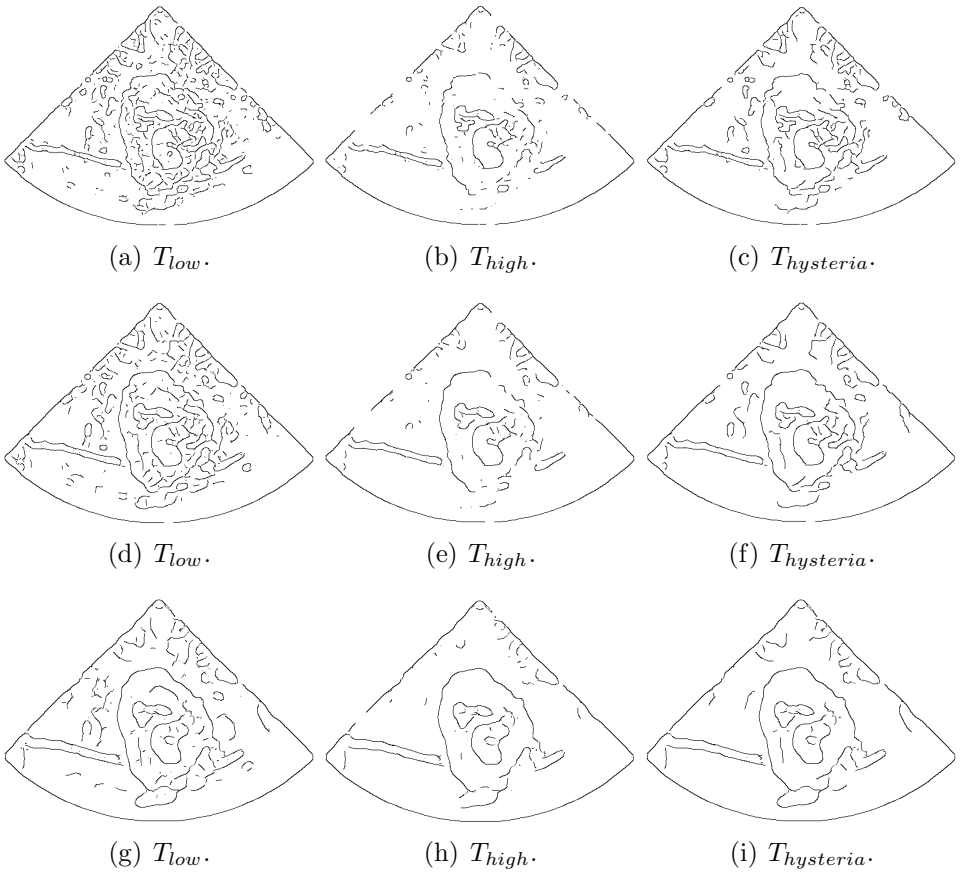
[1] Canny argued in [3] that there is typically little to gain by using several scales. This coincides with the observations made in the 1-D case.

formulated as;

1. Compute $W^x f(u,s)$ and $W^y f(u,s)$ for some $s \in]0,\infty[$.
2. Find the modulus-maximum of $Mf(u,s)$ in the direction given by $Af(u,s)$.
3. Estimate the noise in the signal, e.g. by using a Wiener-filter as presented in Sect.(4.3).
4. Decide a threshold for the probability for an edge to correspond to noise, typically $P_{\%} \in [80,100]$. Find the lower bound $A_{\%}$ for the amplitude of $Mf(u,s)$.
5. Calculate the high threshold $T_{\text{high}} = \frac{A_{\%}}{n_0}\sqrt{s}\Sigma(\psi)$, and the low threshold $T_{\text{low}} = \alpha T_{\text{high}}$ for some $\alpha \in]0,1[$.
6. If a modulus-maximum is above T_{high} it should be marked as an edge. If a modulus-maximum is above T_{low} and connected to a line-segment which at some point is above T_{high}, the modulus-maximum should be marked as an edge.

Without using a noise-estimate skip Step 3 and Step 4 and guess appropriate values for T_{high} and T_{low}. Eventually one can skip Step 3 and guess a value for $A_{\%}$ in Step 4. When using thresholding with hysteria one can with advantage use a high value of T_{high} (or $A_{\%}$). Similarly to the 1-D case, the representation obtained by using a noise-estimate typically contains too much unwanted information. In this thesis, $A'_{\%} \approx 2A_{\%}$ has been an appropriate starting point for finding a usable representation. $A_{\%}$ is the threshold obtained by using a noise-estimate. An appropriate starting-point for the low threshold has been $T_{\text{low}} \approx 0.3T_{\text{high}}$.

Fig.(7.1) illustrates the algorithm applied to the US-image in Fig.(1.4). One thing which may be observed, is the relation between SNR and LOC. At small scales the details of the representation are high, but there are a large number of unwanted edges. This corresponds to a high value of LOC and a small value of SNR. At coarse scales the representation of the objects is of low detail, but there are few unwanted edges. This corresponds to a low value of LOC and a high value of SNR, and agrees with the discussion in Sect.(4.2). In Fig.(7.1,b,e,h) one may observe the phenomenon of streaking. This phenomenon is mostly present at small scales. This is



(a) T_{low}. (b) T_{high}. (c) $T_{hysteria}$.

(d) T_{low}. (e) T_{high}. (f) $T_{hysteria}$.

(g) T_{low}. (h) T_{high}. (i) $T_{hysteria}$.

Figure 7.1: Edge-detection using thresholding with hysteria. The representations are found with respect to the "optimal" threshold, i.e. noise-estimation has not used. (a)-(b)-(c) $s = 3$. (d)-(e)-(f) $s = 4$. (g)-(h)-(i) $s = 6$.

not surprising taken into account the relation between SNR and the scale. At coarse scales it should be easier to separate the modulus-maximum corresponding to important objects, such that the entire line-segment is preserved. Another observation regarding streaking is that it in particular occurs at the upper-right of the tumour and in the region marked **C** in Fig.(1.5). In these regions the transition between tissues changes slowly, i.e. the edge is smooth. This causes the value of the wavelet-transform in these regions to be smaller than in regions where the transition between tissues is sharp. One may for instance observe that streaking is not a problem for the remainder of the

boarder marked **A** and for the cysts.

The advantage of this method is its simplicity, and that it requires little computational effort. In addition it is a nice property that a threshold can be pre-estimated. In this case the edge-detector requires no feedback from an operator. Unfortunately, the representation obtained by using a pre-estimated threshold, e.g. by using Wiener-filtering, is not adequate with respect to the purpose of this thesis. Better representations are obtained by a manual investigation with different values for the threshold. The disadvantage is that it may require a lot of effort to find the threshold which gives the best representation of the signal. Similarly as in 1-D, one will have to change both the threshold and the scale in order to find the best representation.

To finish the discussion, the representations obtained in Fig.(7.1,c,f,i) are compared with respect to the preliminary targets defined in Sect.(1.2.2). All the representations fulfil the first minimum-criteria; to detect the contours marked **A**,**B** and (something) like the yellow-dotted line. The representations more or less fail to fulfil the second criteria; all representations have additional unwanted edges. For the representation in Fig.(7.1,c,f) this is unsatisfactory. For the representation in Fig.(7.1,i) the number of additional edges are low. This representation is acceptable. With respect to the optional requirements discussed in Sect.(1.2.2), the representations in Fig.(7.1,c,f) give a high detail representation of the upper-right cyst, and detects the objects marked **D**. The last representation does not fulfil any of the optional requirements.

The conclusion regarding this edge-detectors performance to noisy images, is similar as the for the 1-D edge-detector. With respect to find an representation of the image, there is a pay-off between details and the reliability. At fine scales one can detect small objects, but there is a large presence of unwanted information. In addition one may have problems to detect an edge representing a slow change of state in the signal, i.e. a smooth edge. At coarse scales one will avoid the presence of unwanted information and detect smooth edges, but not to detect small objects in the image.

7.1.2 Spatial Edge-Detector.

In this section a single-scale edge-detector exploiting the length and the amplitude of a line-segment will be discussed.

In the previous section it was concluded that using information of the amplitude of a single modulus-maximum, was not appropriate in order to obtain a nice representation of the US-image. In Sect.(5.2.2) it was discussed

that a line-segment corresponding to a major object, typically is longer than for noise and small objects. The edge-detector which will be discussed in this section, combines this with the observation that a large edge has a corresponding large value of the modulus-maximum. A nice property of this edge-detector, is that it can easily be combined with a multi-scale edge-detector which will be presented in Sect.(7.2.2). The resulting edge-detector will prove effective for analysing the US-image. In Sect.(6.2) it was discussed that one disadvantage with using P_1 to find the edges, was that it was difficult to pre-estimate or guess an appropriate threshold. In this section it will be discussed how one in certain cases can obtain an automatic threshold for P_1 and similar quantities which will be introduced in this and the following section.

Two slightly different methods will be presented in this section; with or without connecting line-segments inside a scale. The motivation, advantages and how one can connect line-segments inside a scale, was discussed in Sect.(5.2.2).

Based on the output of the detector it will be observed that this edge-detector works well to detect large objects with a sharp contour. It has problems to detect small objects, and to detect objects with a vague contour. The effect of connecting line-segments inside a scale will be illustrated.

The section begins with a discussion of the underlying idea behind the edge-detector, and how a threshold can be found. The algorithm will be presented and applied to the image in Fig.(1.4).

The underlying idea behind this edge-detector is based on two observations; a large step-edge has corresponding large value of the modulus-maximum, and a line-segment corresponding to a major object is typically longer than for noise etc. These observations are supported by Fig.(3.3), Fig.(3.4) and the discussion in Sect.(5.2.2). One problem with considering the length of a line-segment is that it tends to break up towards finer scale. This is a disadvantage if one tries to exploit the size or circumference of an object to detect the major objects. In Sect.(5.2.2) it was discussed how one could connect line-segments inside a scale, and therefore preserve information about circumference or size of the object. The edge-detector which is discussed in this section can be applied both with and without connecting line-segments inside a scale. Note that if one uses connected line-segments, the algorithm is no longer a single-scale edge-detector (because it uses information from several scales).

To each (class of connected) line-segment $\{m_i\}_{i=1}^{N}$ where N is the number of distinct (classes of connected) line-segments, the following quantity is assigned;

$$P_{\text{sp}}(m_i) = L_i \cdot \bar{A}_i.$$

L_i is the length of the (class of connected) line-segment m_i, and $\bar{A}_i$ is the average amplitude of the (class of connected) line-segment corresponding to m_i. After this quantity is assigned to each (class of connected) line-segment, one has to find a threshold separating the values corresponding to noise and important edges. The following paragraph discusses how a threshold can be obtained, and introduces the *natural threshold*.

In Sect.(6.2) it was discussed that one disadvantage with using P_1 to detect the edges in the 1-D signal, was that one could not guess or estimate a threshold. In this paragraph it will be discussed how one sometimes can find such a threshold, denoted the *natural threshold*. Whether such a threshold exists or not, depends on the quantity used to represent the line-segments/modulus-maximum in the signal. The following method has been very successful for the US-image in Fig.(1.4) with respect to some of the quantities used in this thesis. In order to find a threshold-value, one displays the values of P_{sp} in a histogram. If the signal is sufficiently nice (with respect to P_{sp}), the threshold will appear almost by itself in the histogram. The idea is that the quantity P_{sp} divides the line-segments into two disjoint sets, one assumed to contain all the uninteresting information and the other the interesting information. The following discussion holds only if the quantity P_{sp} yields such a separation of the modulus-maximum. For instance for the quantity P' = (amplitude of a modulus-maximum) (equivalent to the previously discussed detector), one can not find a threshold by the following presented approach for the US-image investigated in this thesis. Fig.(7.2) illustrates what is meant with the statements "threshold-value appears by itself" and two disjoint sets. The first set is assumed to contain the large infrequently occurring values. With respect to P_{sp} this corresponds to line-segments of long length and large average amplitude. The other set is assumed to be the cluster concentrated at small values. This cluster contains line-segments of small length and small average amplitude, such as noise and small objects. The cluster behaviour in Fig.(7.2) is a typical behaviour of noise. The threshold is assumed to be a value separating the two sets, i.e. separate the large cluster from the infrequently occurring values. The term *natural threshold* will refer to the smallest value which separates the cluster from the discretely

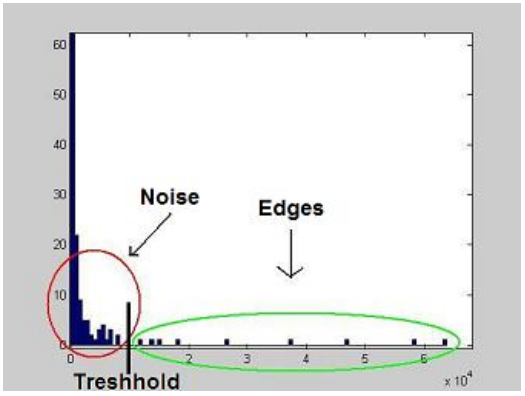


Figure 7.2: Illustration of a histogram illustrating the values of P_{sp} for a signal. Values of P_{sp} corresponding to noise is assumed to be gathered in the cluster at low values. A threshold value is some value separating the cluster from the infrequently occurring values.

distributed values. This threshold will depend on the size of the histogram, so some slack has to be allowed. In this thesis, the histogram-size has been chosen to be 100 for all the quantities P_1, P_{sp}, P_{sc} and P. As discussed in the previous chapter, one disadvantage with using P_1 was the difficulty of guessing or estimate a threshold which would give a usable representation of the signal. By using the natural threshold one can obtain a usable representation. The ideal situation is if an almost optimal representation follows from the natural threshold. This will depend on how well the quantity used to represent the line-segments/modulus-maximum is adapted to the signal with respect to which information is wanted. In this and the following section, it will be discussed how the quantities P_{sp}, P_{sc} and P is adapted to the US-image in Fig.(1.4). The latter two quantities will be introduced in the next section. This method for finding the threshold may appear unreliable and random. In fact this method has been used with success for the majority of the representations in this chapter and in the appendix.

The algorithm can be formulated as;

1. Compute $W^x f(u,s)$ and $W^y f(u,s)$ for some $s \in]0,\infty[$. Find the modulus-maximum of $Mf(u,s)$ in the direction given by $Af(u,s)$.

2. Identify the line-segments $\{m_i\}_{i=1}^N$. N is the number of line-segments

at scale s.

3. Find the length L_i of each line-segment for $i \in \{1, ..., N\}$.

4. Find the average amplitude $\bar{A}_i$ of each line-segment for $i \in \{1, ..., N\}$.

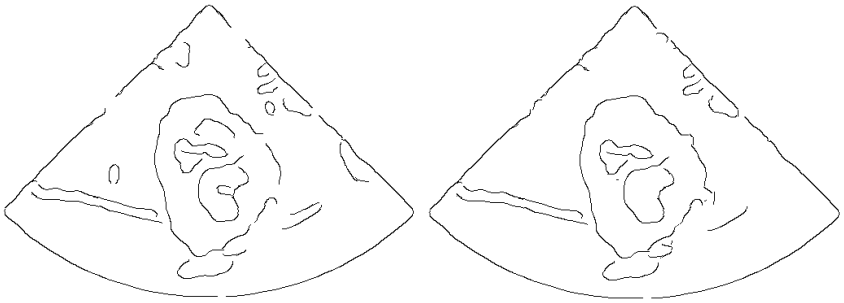
5. Calculate the product $P_{\text{sp}}(m_i) = \bar{A}_i \cdot L_i$, and display the values in a histogram.

6. Find an appropriate threshold, and remove the line-segments with P_{sp} value below the threshold.

If the line-segments have been connected, one uses "class of connected line-segments" rather than "line-segment" in the above algorithm

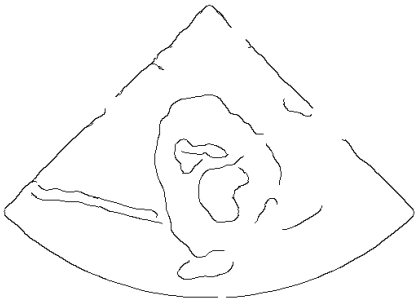
Fig(7.3,a,c) illustrates the output of this algorithm, without connecting line-segments inside a scale. One may observe that a problem occurs in the regions where the contour of the tumour is vague, e.g. in the region marked **C** in Fig.(1.5). The contour of an object in such regions is often broken into several small line-segments. In addition these edges are smoothened, and as discussed in Sect.(2.6), the modulus-maximum corresponding to a smooth edge decay faster towards finer scales. The value of P_{sp} of these line-segments will often be hidden in the cluster making them difficult to recover without including unwanted information. Observe that the representation of the cysts and **A** are equal in both the algorithm with and without connecting line-segments. The contour for these objects is sharp.

In Fig.(7.3,b,d) the line-segments are connected by the method discussed in Sect.(5.2.2). Two "false" connections appear in the output. The upper line of the cortex is considered to be connected to the boarder of the US-image, and the line-segments underneath the tumour are considered to be one line-segment. Both of these "false" connections are caused by the phenomenon discussed with respect to Fig.(5.8) in Sect.(5.2.2). The reason for the false connection, is found by considering Fig.(3.3). One may observe the increased performance in the low-contrast regions, compared to the representations in Fig.(7.3,a,c). Because this edge-detector exploits information from coarse scales, one will avoid problems caused by breaking-up of line-segments. One may observe that there are no small gaps in the representation, e.g. to the upper-right of the tumour and upper left of the boarder of the image. This helps to obtain a nice representation of the image. Note that the natural threshold in this case gives an almost optimal representation.



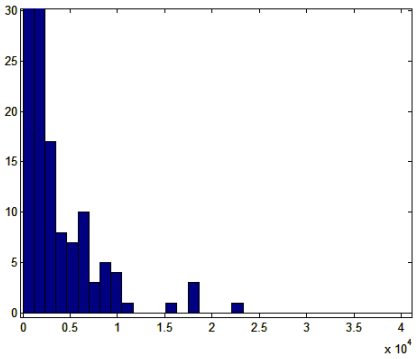
(a) $T_{P_{\rm sp}} = 7500$.

(b) $T_{P_{\rm sp}} = 15000$.

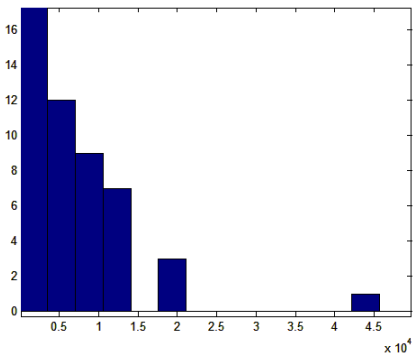


(c) $T_{P_{\rm sp}} = 15000$.

(d) $T_{P_{\rm sp}} = 25000$.

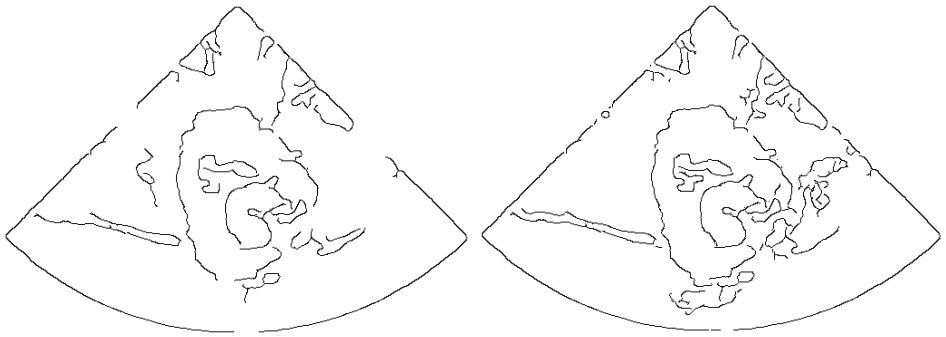


(e) Histogram of the values of $P_{\rm sp}$ without connecting line-segments inside scales.

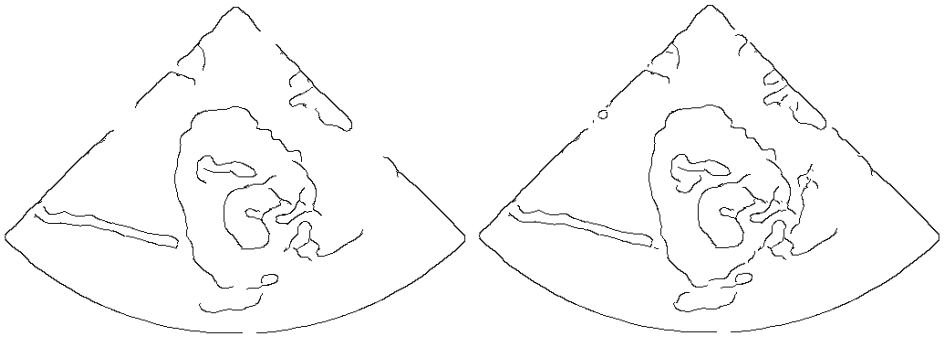


(f) Histogram of the values of $P_{\rm sp}$ with connecting line-segments inside scales.

Figure 7.3: The edges detected by using the quantity $P_{\rm sp}$ at $s = 5$. (e) - (f) The histogram is enlargened in the region of the separation between the "noise-cluster" and the infrequently occuring values.



(a) s=3. (b) s=3.



(c) s=4. (d) s=4.

Figure 7.4: Illustration of the edges detected using P_{sp}. (a) - (c) No preprocessing has been done for the edges. (b) - (d) The line-segments have been connected inside the scale as described in Sect.(5.2.2).

One advantage of this algorithm is its ability to remove small line-segments corresponding to noise. One may observe that there are only a few unwanted edges in the representations. Another advantage of this method, is that it preserves connectedness of line-segments. This is desired in order to obtain a nice representation of the signal. A disadvantage is that objects with small circumference maymbe hidden among the noise. For instance it may be troublesome to detect the objects marked **D** in Fig.(1.4). This may be observed in Fig.(7.3a), where one of the objects marked **D** is recovered by including a part of the cluster at small values.

Fig.(7.4) illustrates the representations obtained at fine scales, with and without connecting line-segments inside scale. One may wonder why one of

the small objects marked **D** in Fig.(1.5) are present in all representations. This is caused by the simple reason that this object at these scales is connected to the boarder of the large cyst. In Fig.(7.4c) one may observe that a piece of the contour of the upper cyst is removed. The line-segment representing this contour is disconnected at $s = 4$. One of these line-segments are not sufficiently long in order to not being discriminated by P_{sp}. Compared with the fine scale representations in Fig.(7.1,c,f) one may observe that P_{sp} is good for removing unwanted information, but struggles in low-contrast regions. When using P_{sp} with connecting line-segments inside the scale, one may observe increased performance in the low-contrast regions. One has some problems with removing unwanted information to the right of the tumour. The line-segments at the upper-right are at these scales connected to the boarder of the image, and are not possible to remove.

The representation in Fig.(7.3c) and Fig.(7.4,a,c) fails to fulfil the minimum-requirements defined in Sect.(1.2.2), and do not fulfil any of the additional requirements. The representations in Fig.(7.3,a,b,d) and Fig.(7.4,b,d) do fulfil the minimum-requirements. They give a nice representation of the contours marked **A**, **B** and the yellow-dotted line in Fig.(1.5).

The edge-detector which use P_{sp} works well for detecting objects of large strength and large circumference in the image, but will not detect small objects. At fine scales P_{sp} will not detect an edge corresponding to a vague contour, i.e. a smoothened edge.

7.2 Multi-Scale Edge-Detectors.

7.2.1 Lipschitz Edge-Detector.

This section will present a 2-D multi-scale edge-detector based on the Lipschitz-regularity of the modulus-maximum.

In Sect.(3.3) the relation between the decay of the CWT and Lipschitz-regularity was established. This relation enables one to distinguish different kind of edges in the signal, such as noise and step-edges.

In Sect.(5.2) it was discussed how modulus-maximum could be connected across scales. The 2-D maxima-tree plays the main role for the edge-detector presented in this section. The Lipschitz-regularity is invariant of the strength of an edge, i.e. the Lipschitz-regularity does not tell which edge is larger than the other. In order to find which edges are more important, one can use the

length of the maxima-line in time-scale plane.

This algorithm will be applied to the image in Fig.(1.4), and the performance will be discussed. It will be verified that the edge-detector is acceptable for detecting the main-characteristics in the image. One characteristic feature of this method is its "invariance" to the strength of the edge. It will be concluded that thresholding the length of the maxima-lines is not sufficient to remove edges of small strength. Another characteristic feature of this edge-detector is its ability to remove edges of short duration.

The section will begin with a discussion of the uncertainty of how a threshold of the Lipschitz-regularity can be found, before presenting the algorithm. At the end this edge-detector will be applied to the image in Fig.(1.4) and the result discussed.

The 2-D algorithm which separates modulus-maximum based on its Lipschitz-regularity and the length of the maxima-line, is identical to the algorithm presented in Sect.(6.2). Similarly as in 1-D, one problem is to decide a threshold for the Lipschitz-regularity. Consider the histogram in Fig.(7.5) illustrating the average Lipschitz-regularity of the modulus-maximum at scale $s = 5$. One may observe that there is no separation of the values of the

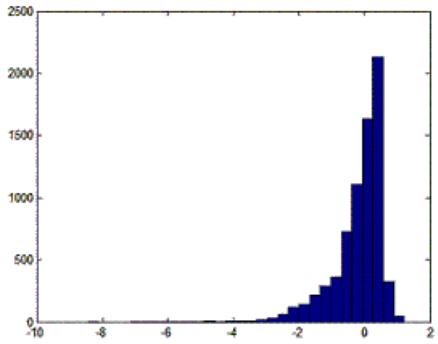
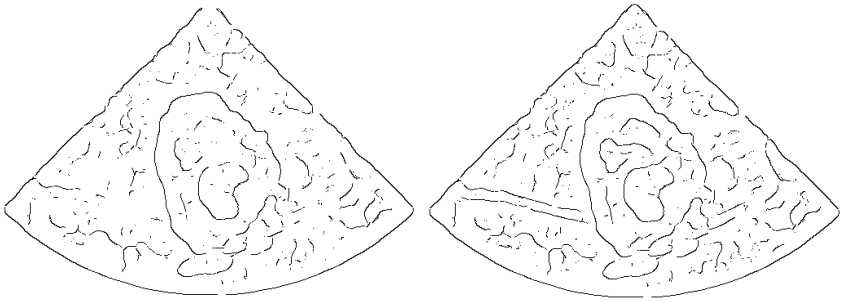


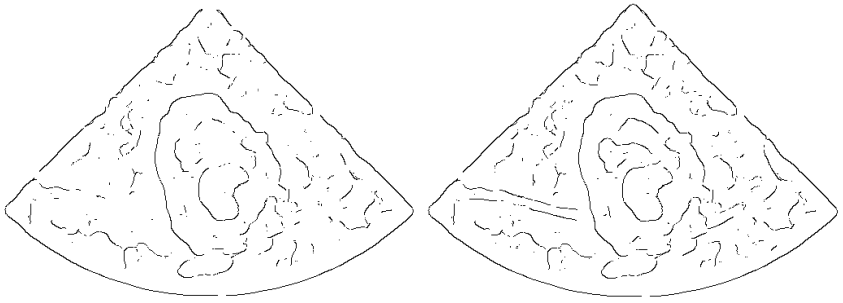
Figure 7.5: The average Lipschitz-regularity of the maxima-lines converging to the modulus-maximum at scale $s = 5$.

Lipschitz-regularity. The "connectedness" of the Lipschitz-values is a problem for estimating a threshold for the Lipschitz-regularity.

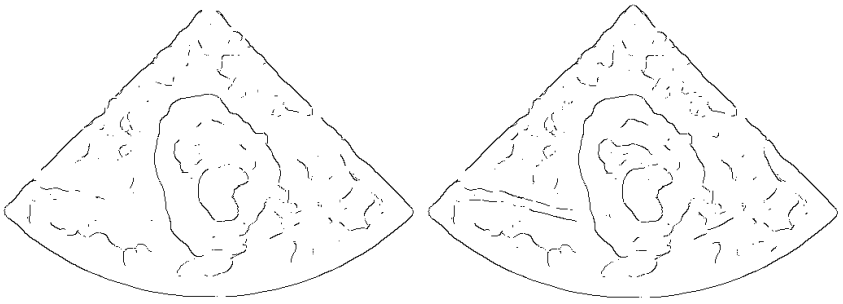
Fig.(7.6) illustrates this edge-detector applied to the image in Fig.(1.4). One may observe that thin edges are discriminated. This may be observed



(a) $T_{Lip} = -0.2$, $L = 0$. (b) $T_{Lip} = -0.6$, $L = 0$.



(c) $T_{Lip} = -0.2$, $L = 3$. (d) $T_{Lip} = -0.6$, $L = 3$.



(e) $T_{Lip} = -0.2$, $L = 5$. (f) $T_{Lip} = -0.6$, $L = 5$.

Figure 7.6: The edges detected by using the average Lipschitz-regularity and the length of the maxima-lines.

in the representation of the upper cysts, the separation between the cysts and the cortex. These edges correspond to a change of state in the signal of short duration. Their corresponding Lipschitz-estimate are therefore low. Unfortunately this property is not required for the applications of this thesis.

The second characteristic-feature of the algorithm is its "invariance" to the strength of the edge. From Fig.(7.6) one may observe that there are several responses which are not removed by thresholding the length of the maxima-lines. One example is the small line-segment above the cyst at the top. The edge this line-segment reflects, is not of large strength, see Fig.(1.4). An interesting observation is that the more strength of an edge, the longer spatial distance before small edges is included. This is in particular evident around the contours marked **A** and **B** in Fig.(1.5). One may observe the buffer-zone around these edges where there are none of the small edges. This follows (almost) by the discussion in Sect.(5.1). The distance to parallel line-segments are forced to be large by a similar reason as discussed in Sect.(5.1). To perpendicular line-segments, the distance can be closer. In these regions thresholding the length of the maxima-lines is be effective. In regions where the distance to strong edges is large, this effect decreases. As in 1-D, this may be adjusted by including coarser scales in the 2-D maxima-tree.

With respect to the preliminary-targets defined in Sect.(1.2.2), the edge-detector satisfies the first target; to find the contour of the tumour. The performance to detect the cysts is not satisfactory. Taken into account the edge-detectors discussed previously, this edge-detector performs badly in preserving connectedness in the representaion, and it contains a lot of unwanted information.

The conclusion is that the Lipshitz edge-detector works well for finding the points where the signal changes state with a long duration, and is effective for removing edges of short duration.

7.2.2 Scaling Edge-Detector.

This section will present an edge-detector which uses the amplitude of line-segments across scales to find the important edges in a signal.

The quantity P_{sc} which will be introduced in this section, is similar to P_1 discussed in Sect.(6.2). The difference is that instead of considering the amplitude of a modulus-maximum across scales, one will in this section consider the amplitude of a line-segment across scales. In Sect.(7.1.2) it was concluded that P_{sp} had difficulties to detect edges in regions with low contrast. In this section it will be investigated if one can detect edges corresponding to objects with vague contours by using the evolution of line-segments across scales. In addition it will be investigated if this edge-detector detects objects of small size in the image. The evolution of line-segments across scales contains useful

information which can be used to find the important edges in a noisy-image. This was observed for the similar 1-D edge-detector in Sect.(6.2).

In order to investigate whether P_{sc} works better to detect vague contours and small objects in the US-image in Fig.(1.4) one has to be careful. In order to investigate how it handles vague contours, it will not make sense to use line-segments which are connected inside scales. As one may observe in Fig.(7.3) one effectively avoids the problem with the vague contour underneath the tumour by connecting line-segments inside scales. Therefore this problem will be investigated by using an alternative approach for connecting line-segments.

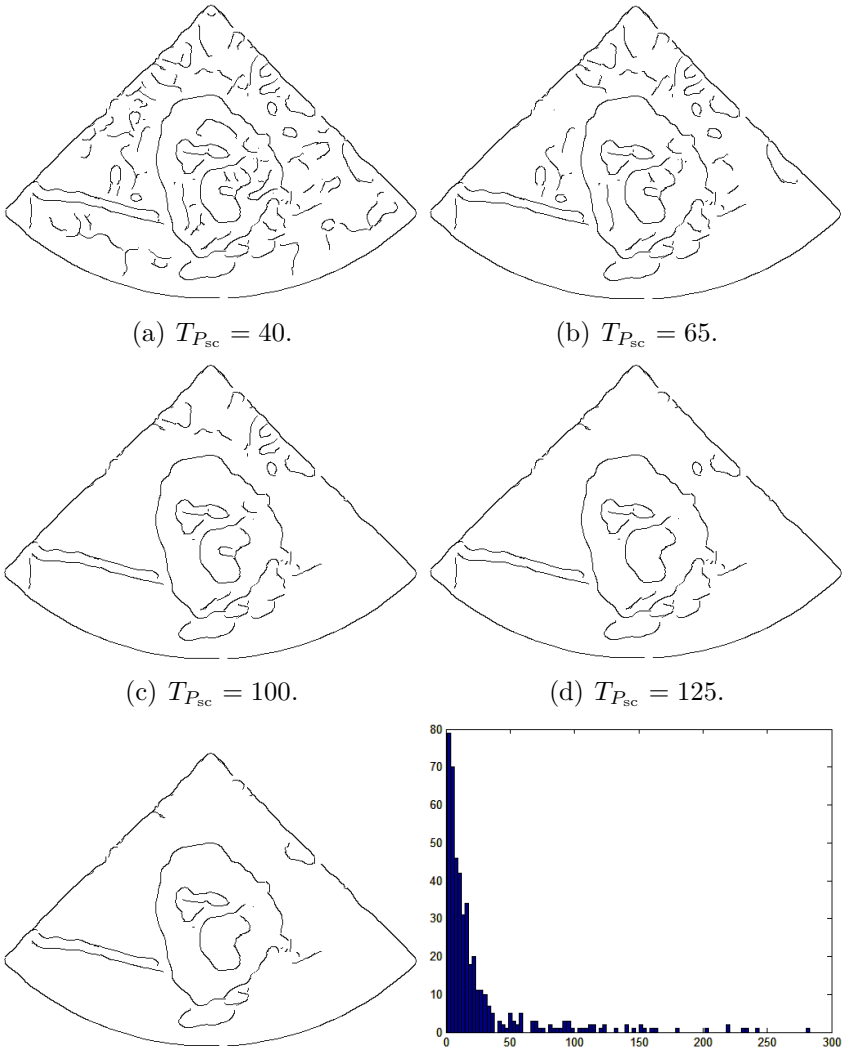
Based on the output of the edge-detector it will be concluded that the quantity P_{sc} works well in to detect vague contours and to detect small objects. A disadvantage with P_{sc} is that it is less useful to remove edges of small strength. This is caused by a similar phenomenon as for the Lipschitz edge-detector. A line-segment corresponding to a small edge may exist for several scales. In addition a major obstacle concerning convergence of line-segments across scales will be visualised.

The section begins with a presentation of the algorithm for the edge-detector, and discussing an alternative method for connecting line-segments across scales. This algorithm will be applied to the image in Fig.(1.4). The output will be discussed with respect to the preliminary targets defined in Sect.(1.2.2). At the end of the section, the edge-detectors in this section and Sect.(7.1.2) will be combined.

In Sect.(5.2.2) it was discussed how one could relate line-segments across scales and obtain an "1-1" relation between line-segments across scales. The motivation was to ensure that the objects used in the computations were similar with respect to the length. This was done by using information of the line-segment at coarse scale, and try to make the line-segments at finer scale to inherit this information. Assume there are N line-segments $\{m_i\}_{i=1}^N$ at the finest scale, and let $[m_i]$ denote the line-segments which converge towards m_i. Assign to each line-segment at the finest scale the quantity;

$$P_{\text{sc}}(m_i) = l_i \cdot \bar{a}_i.$$

l_i is the length in time-scale plane of the line-segment which converge towards m_i, i.e. the number of scales there exist line-segments converging towards m_i. $\bar{a}_i$ is the (weighted) average amplitude of the line-segments which converge towards m_i. The algorithm can be summarized as:



(a) $T_{P_{\text{sc}}} = 40$. (b) $T_{P_{\text{sc}}} = 65$.

(c) $T_{P_{\text{sc}}} = 100$. (d) $T_{P_{\text{sc}}} = 125$.

(e) $T_{P_{\text{sc}}} = 170$. (f) Histogram of the values of P_{sc}.

Figure 7.7: The edges detected by applying P_{sc} to line-segments. The line-segments have been connected inside scales. The scales used in the computation are $s \in [5, 10]$.

1. Compute $W^x f(u,s)$, $W^y f(u,s)$ and $Mf(u,s)$ for $s \in \{s_1, ..., s_J\}$.

2. Compute $Af(u,s)$, and find the modulus-maximum of $Mf(u,s)$ in the

direction given by $Af(u,s)$ for $s \in \{s_1, ..., s_J\}$.

3. For each line-segment m_i at scale s_1, find the set $[m_i]$ of line-segments at coarse scales which converge towards m_i.

4. Assign to each m_i, the quantity $P_{\text{sc}}(m_i) = l_i \cdot \bar{a}_i$.

5. Find a threshold which separates the noise from the important edges.

An alternative method to ensure that line-segments across scale are similar with respect to the length, is to divide the line-segments at coarse scale. Instead of using a line-segment at coarse scale to connect line-segments at finer scales, one can use the line-segments at fine scale to divide line-segments at coarser scales. How this can be done in practice will not be discussed in this thesis. Instead a short cut has been used. The information which is wanted to obtain from coarse scales, is the average amplitude of the (disconnected) line-segment which converges towards the line-segment at finest scale. Let m_i denote a line-segment at the finest scale. One needs an estimate of the average amplitude of the line-segment at scale s which converges towards m_i. This estimate is defined to be the (weighted) average amplitude of the modulus-maximum at scale s which converges towards m_i. In Sect.(5.2.1) it was discussed how modulus-maximum converges in 2-D. At the end of this section the above algorithm will be used with disconnected line-segments rather than connected line-segments, and explained why we are interested in this.

Fig.(7.7) illustrates the output of the "first" algorithm applied to the image in Fig.(1.4). I.e. line-segments have been connected inside scales. The "false" connections are the same as those discussed in Sect.(7.1.2). One may observe in the histogram, that the separation between the cluster and the infrequently occurring values is less evident than for P_{sp}. It is clear that there is a cluster of noise at small values, and a cluster of infrequently occurrences at high values, but the exact point of separation is vague. Both 40, 65, 100 and 125 are justifiable as the natural threshold for P_{sc}. Note that this threshold will depend on the histogram-size. One reason why there is not an equally evident separation as for P_{sp}, is that P_{sc} does not use information about the length of the line-segments. In an image with objects of large circumference, the P_{sp}-values of these objects become very large. One problem with P_{sc} is that it may include edges of small strength. This is caused by the similar phenomenon which caused difficulties for the Lipschitz

algorithm. An edge of small strength may exist for several scales if the spatial distance to edges of large strength is sufficiently long. One evident example is the line-segment underneath the large cyst in Fig.(7.7,a,b,c,d). The edge which this line-segment represents is not of large strength. One advantage with P_{sc}, is that it does not discriminate edges based on its size, i.e. if its line-segment is short. For instance one of the objects marked **D** in Fig.(1.5) is present in Fig.(7.7,a,b,c).

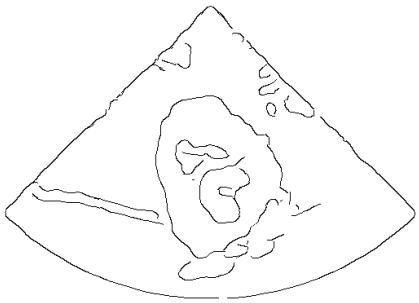
The output satisfies the minimum-requirements for the edge-detector. The contour of the tumour and the cysts are all present. For high values of the threshold $T_{P_{\text{sc}}}$ there are few unwanted edges. Observe that the contour in the lower half is nice at all scales. For the representation in Fig.(7.7a,b,c) one of the objects marked **D** in Fig.(1.5) is represented.

One nice property of this edge-detector is that it can easily be combined with the edge-detector presented in Sect.(7.1.2). To each line-segment at the finest scale assign the quantity;

$$P(m_i) = P_{sp}(m_i) \cdot P_{\text{sc}}(m_i).$$

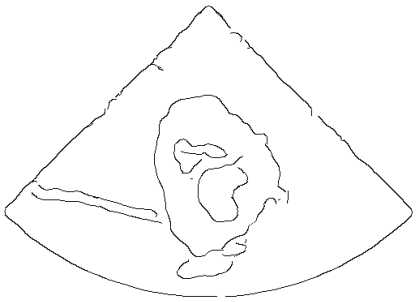
Consider Fig.(7.8) illustrating the output with respect to P. One may observe that some of the unwanted edges which has been present after using P_{sp} and P_{sc} separately, are removed in this representation. The representation obtained in Fig.(7.8c) is the superior representation with respect to the preliminary targets defined in Sect.(1.2.2). Note that the natural threshold gives an almost optimal representation.

One question is still not answered. The purpose of introducing P_{sc} was to improve the performance to detect vague contours, i.e. smooth edges. When using line-segments which have been connected inside scales, one effectively avoids the problem-area underneath the tumour for the US-image in Fig.(1.4). This can be observed in Fig.(7.3,b,d). In order to verify whether the performance of P_{sc} (and P) increases in regions with low contrast, the algorithm will be applied for line-segments which have been disconnected. I.e. the line-segments have been divided towards coarse scales rather than connected towards finer scale. In Fig.(7.3,a,c) one may observe that in this case P_{sp} struggles underneath the tumour. The target is to see if P_{sc} and P performs better to detect the contour under the boarder. Fig.(7.9) illustrates the output of P_{sc} and P with respect to the second approach. There are some problems in representations in (a) and (c) under the tumour. This result is

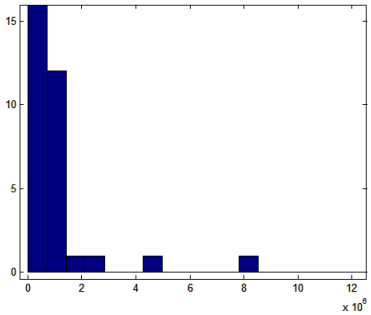


(a) $T_P = 1e6$.

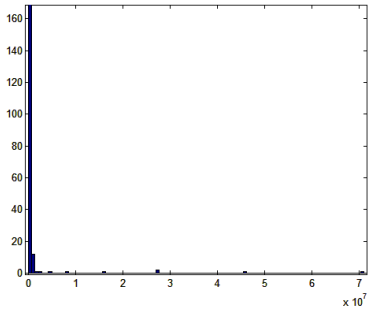
(b) $T_P = 1.5e6$.



(c) $T_P = 6e6$.



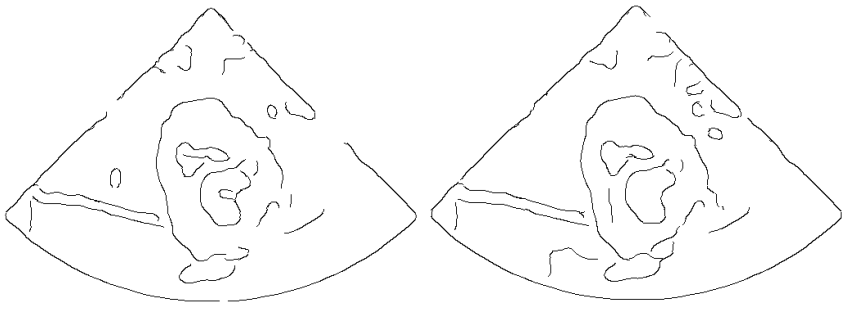
(d) Histogram of the values of T_P.



(e) Histogram of the values of T_P.

Figure 7.8: The edges detected by using P. (d) - The histogram is enlarged in the region of the separation between the cluster and the infrequently occurring values. (e) - The histogram.

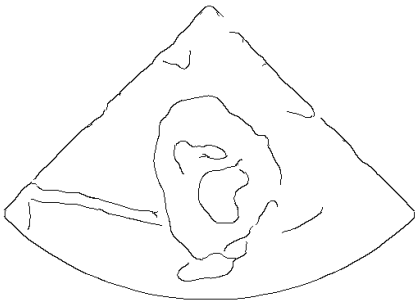
not what one would hope for or expect by using information from several



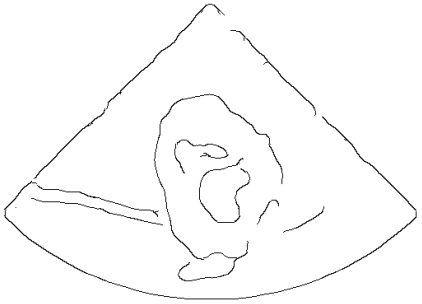
(a) P_{sc}

(b) P_{sc}

(c) P



(d) P



(e) P_{sp}

Figure 7.9: The edges detected by using P and P_{sc}. In this case the line-segments have been disconnected at coarse scales. (a)-(c) $s \in [5, 10]$, (b)-(d) $s \in [6, 10]$. (a)-(b) P_{sc} has been used to find the edges. (c)-(d) P has been used to find the edges. (e) P_{sp} has been used to find the edges at $s = 6$.

scales. What goes wrong is in fact not a weakness of P_{sc}, but a problem with convergence of modulus-maximum in 2-D. One may observe in Fig.(3.3) that

the line-segment representing the boarder underneath the tumour, drastically changes path between $s = 5$ and $s = 8$. The path include the entire region underneath the tumour at coarse scales, and at fine scales changes to a path which is closer to the tumour. This causes that the line-segment "missing" in the representation to exist only for a few scales. The information at coarse scales is in fact gathered in the short line-segment which is underneath the tumour. If the boarder had not changed so dramatically and so fast, it is very likely that P_{sc} would have detected the "missing" line-segment. This statement is supported in Fig.(7.9,b,d,e). In this case, P_{sc} has been computed for $s \in [6, 10]$. At these scales one will avoid the problem of convergence underneath the tumour, see Fig.(3.3). P_{sp} has still problems to detect the contour in this region. One may observe that the representations in Fig.(7.9,b,d) is good with respect to the preliminary targets discussed in Sect.(1.4).

Even with the problems, it is not false to conclude that P_{sc} works better in regions of low contrast. This is supported by the representations of other medical US-images displayed in App.(A).

The problem of convergence of modulus-maximum illustrates one advantage of connecting line-segments inside scales. In this case one will be less vulnerable for such problems. Recall that by relating line-segments from coarse scale, both paths underneath the tumour were used to represent the boarder under the tumour. Another advantage with connecting line-segments can be observed in Fig.(7.9d,e). The representation of the upper cyst is composed of two line-segments. In this case there is only a 1-pixel gap between the line-segments. The length of the smallest line-segment is not sufficient to prevent it from being discriminated by P and P_{sp}.

The conclusion is that P_{sc} works well to detect edges corresponding to vague contours and detect small objects in a noisy image. In addition P_{sc} can detect edges of relatively small strength compared to the strength of the noise. The edge-detector which uses P works well to detect the objects of large strength, and will detect vague contours better than P_{sp}. In addition P will avoid including objects of small strength such as P_{sc}. Each of P_{sc} and P_{sp} weaknesses are the strength of the other. Combined the strengths seem to cancel the weaknesses, resulting in an edge-detector which is good for detecting objects of large strength in the signal.

7.3 Summary.

In the previous two chapters several edge-detectors have been discussed and applied to the image in Fig.(1.4). Some of their most characteristic-features with respect to the analysis of noisy signal are summarized below.

For the edge-detectors which use a threshold of the amplitude of the modulus-maximum at a single scale, there is a pay-off between accuracy and reliability. At fine scales one is able to detect objects of small size, but there is a lot of unwanted information. In addition one may not detect smooth edges. At coarse scales it is easier to detect smooth edges, and there will be less unwanted information in the representation. The disadvantage with using the wavelet-transform at coarse scales is that one may not detect small objects in the signal.

The edge-detector which uses the Lipschitz-regularity of a modulus-maximum and the length of its maxima-line is good for detecting an edge of long duration. An edge of short duration will not be detected. This edge-detector is not suitable to detect small objects in the signal, but is suitable to find points where the signal changes of small strength over a long duration.

P_{sp} works well for detecting objects with large circumference and large strength, but may not detect small objects. At fine scales this edge-detector may have difficulties to detect an edge representing a vague contour of an object in the image, i.e. a smooth edge.

The edge-detector based on P_{sc} works well to detect small objects in the signal, and detect smooth edges. In addition, this edge-detector may detect edges of a relatively small strength compared to the strength of the noise.

The edge-detector using $P = P_{\mathrm{sp}}P_{\mathrm{sc}}$ inherits the qualities of P_{sp} and P_{sp}. The "weaknesses" of each of the quantities are the others "strengths", and the resulting edge-detector works well in order to find the objects of largest strength in the image.

An experience during the work of this thesis is that it is important to keep control of possible errors and other parameters which would change the result.

7.3.1 Sources of errors.

One source of error is the uncertainty regarding the convergence of modulus-maximum across scales, and how to connect line-segments across and inside scales in 2-D. These uncertainties were discussed in Sect.(5.2.1) and

Sect.(5.2.2). In the latter chapter some faults of the convergence of modulus-maximum and line-segments have been illustrated.

A parameter which has had a great influence on the results, the output, and the discussion regarding the differences between the edge-detectors, are the scales used in the analysis. If one for instance had used $s = 10$ as the finest scale, it would be very little that would differ between the edge-detectors with respect to the analysis of the US-image in Fig.(1.4). Another parameter which would change the result is the range of scales used for the analysis. With respect to the analysis of the image in Fig.(1.4), this would in particular influence the performance of the edge-detector using the Lipschitz-regularity. By using coarser scales, one could take advantage of an increased use of the length of the maxima-tree to remove small edges. This was discussed for the 1-D Lipschitz edge-detector.

7.3.2 Conclusion.

The results of the investigation in this thesis will be discussed in two different parts. The analysis of the ultra-sound image in Fig.(1.4) will be used to conclude which edge-detectors discussed in this thesis are most appropriate for such kind of medical US-images. This discussion is supported by representation of several US-images displayed in App.(A). In addition general edge-detection in images will briefly be discussed.

The edge-detectors using a threshold of the amplitude of modulus-maximum and the Lipschitz-regularity are not suitable for edge-detection in general medical US-images. The quantities P_{sp} and P_{sc} and their respective edge-detectors do often return nice representations of US-images, but they are not reliable for general US-images. The edge-detector using the quantity P is the most reliable for the largest class of medical US-signals. The representations obtained by using P are often close to optimal.

In general edge-detection in images relies on how one is able to design quantities which are able to separate the interesting information from the less interesting. These quantities depend on the quality of the image, and which kind of information is desired to obtain. One can for instance use a different weighting on the average amplitude of line-segments across scales, to e.g. reward coarse scale information. One can reward line-segments which form closed curves, decay with a certain rate or reward objects with a certain shape. The choice of quantity should depend on the signal and the information one would like to obtain.

Appendix A

Edge-Detection in Medical Images.

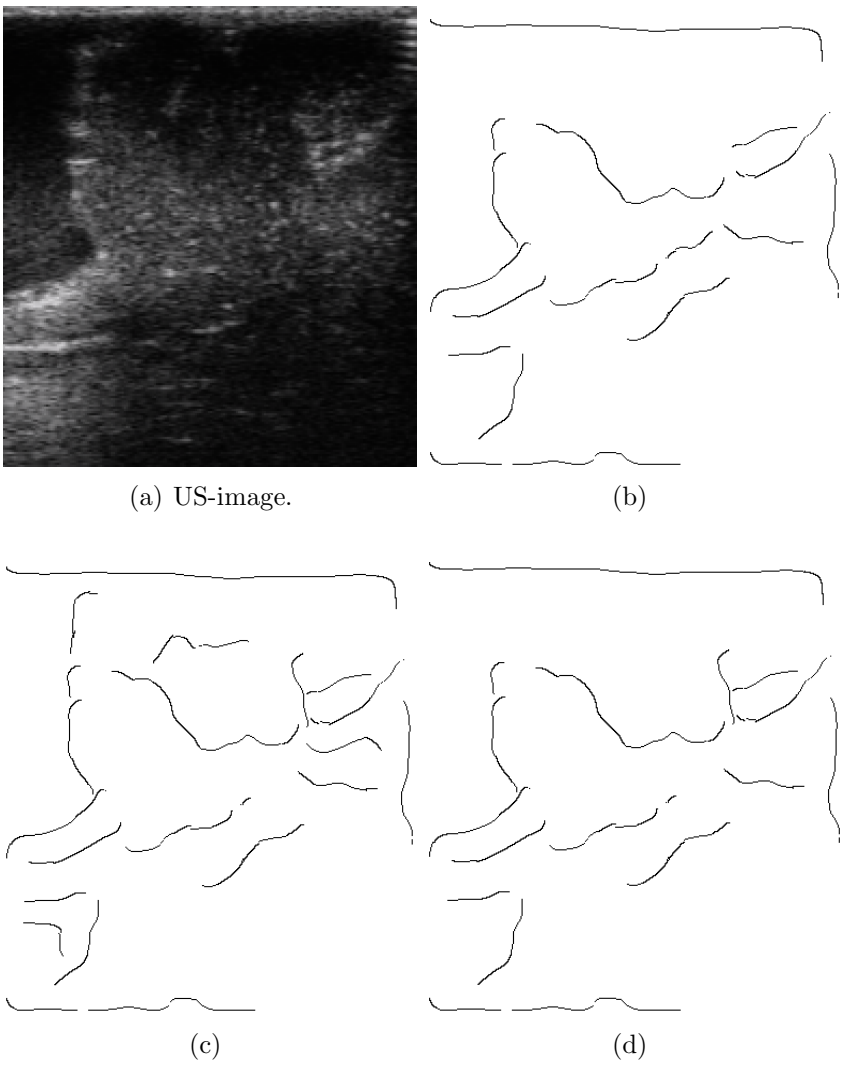
This section contains several representations of medical images using P_{sp}, P_{sc} and P. For each of the images, some observations will be discussed.

A.1 US-image # 1.

The boarder of the object in the US-image in Fig.(A.1a) is vague. In order to obtain a representation, the wavelet-transform has been computed for $s \in [10, 15]$. The interesting information in this US-image is the contour to the middle-left in the image.

A.2 US-image # 2.

For the US-image in Fig.(A.2a) the performance of the edge-detectors were disappointing. For this US-image one should use a larger range of scales. By including coarser scales, one would be able to remove some of the line-segments within the object. One may observe in Fig.(A.2a) that the strength of the boarder at the top of the large object is smaller than for the irregularities within the object. In addition the boarder at the top is vague and indefiniable, causing the representation of the boarder at the top to be broken into small line-segments. This cause that one is not able to recover the contour at the top, without including the irregularities within the object.



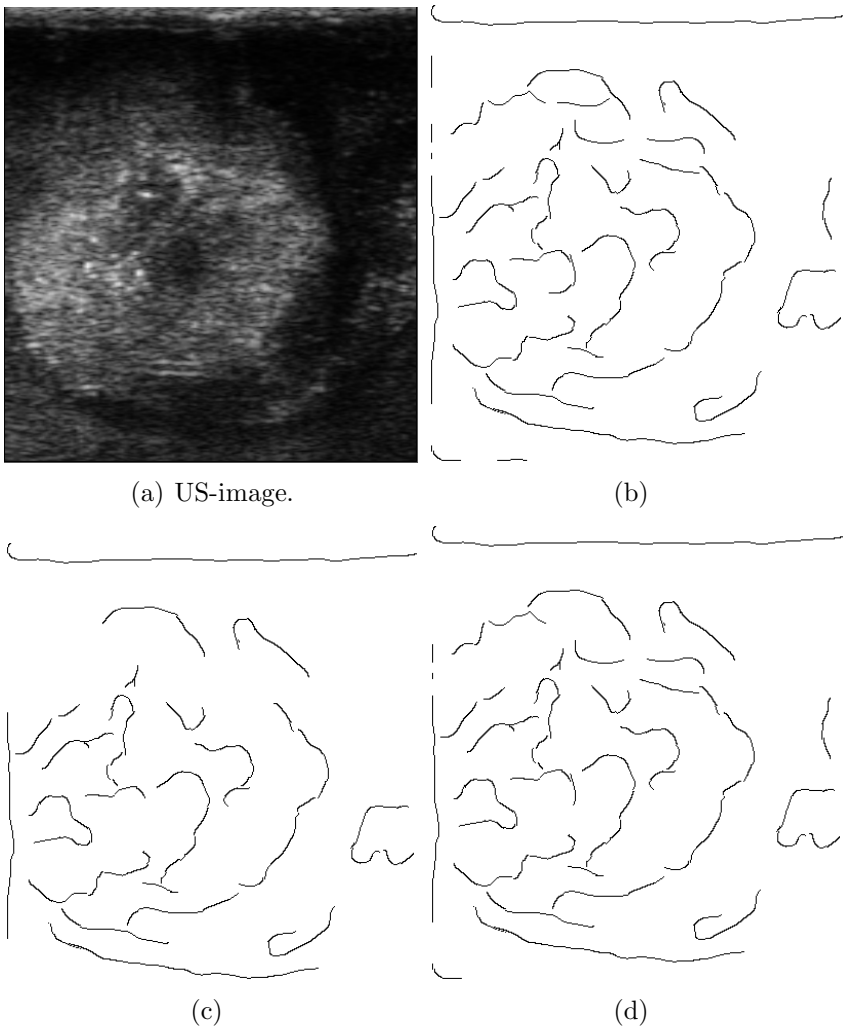
(a) US-image. (b)

(c) (d)

Figure A.1: (a) US-image # 1. (b) The edges detected by using P_{sc}. (c) The edges detected by using P_{sp}. (d) The edges detected by using P. For all the representation a natural threshold has been used.

A.3 US-image # 3.

The representations of the US-image in Fig.(A.3a), is a good example of how P_{sc} and P_{sp} cancel each others weaknesses. One may observe that P_{sp} has

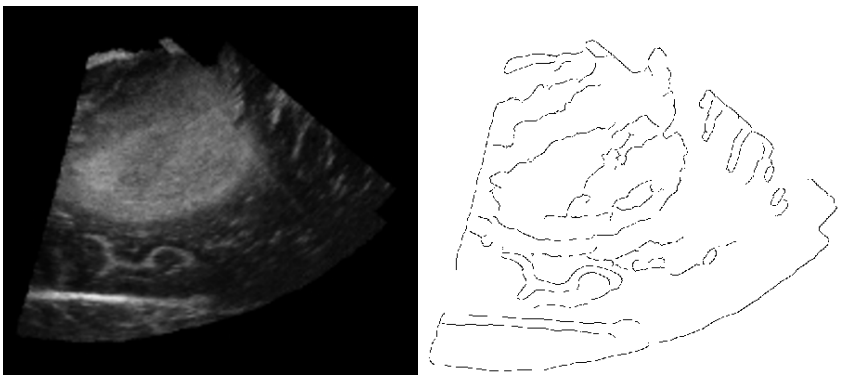


(a) US-image. (b)

(c) (d)

Figure A.2: (a) The US-image. (b) The edges detected by using P_{sc}. (c) The edges detected by using P_{sp}. (d) The edges detected by using P. For all the representation a natural threshold has been used.

problems in the lower right of the large solid white object. The contour of the object in this region is vague. P_{sc} includes a lot of uninteresting information corresponding to objects of small strength. On the other hand P gives a close to optimal representation of the US-image.



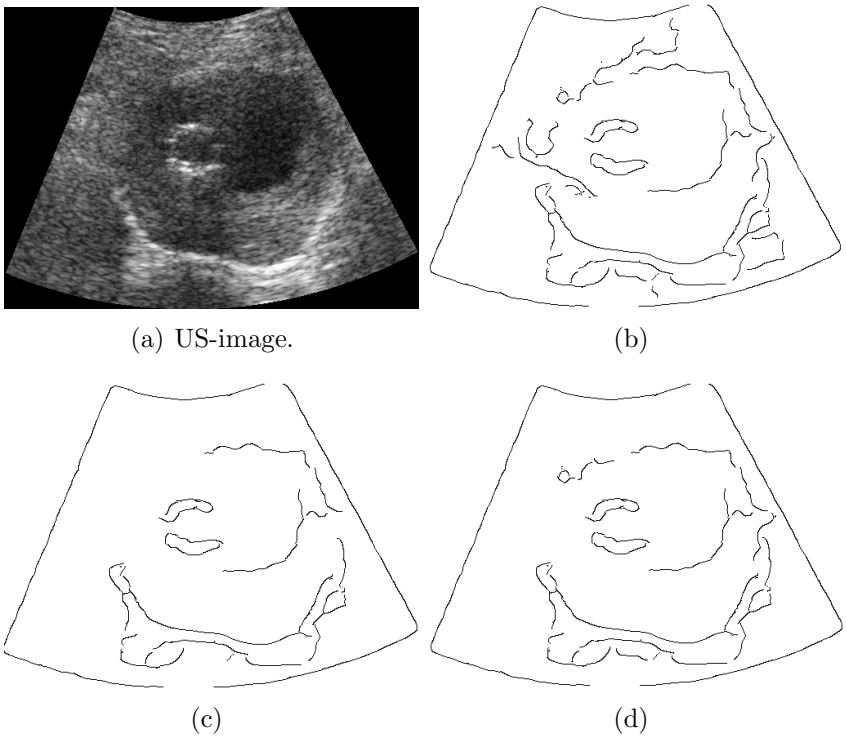
(a) US-image. (b)

(c) (d)

Figure A.3: (a) The US-image. (b) The edges detected by using P_{sc}. (c) The edges detected by using P_{sp}. (d) The edges detected by using P. For all the representation a natural threshold has been used.

A.4 US-image # 4.

The representation of the US-image in Fig.(A.4a) all seem to have a problem with detecting the boarder to the upper-left of the large object. The problem is that this boarder is almost invisible, i.e. the difference between the strength of the object on the inside and the outside of the object is small.



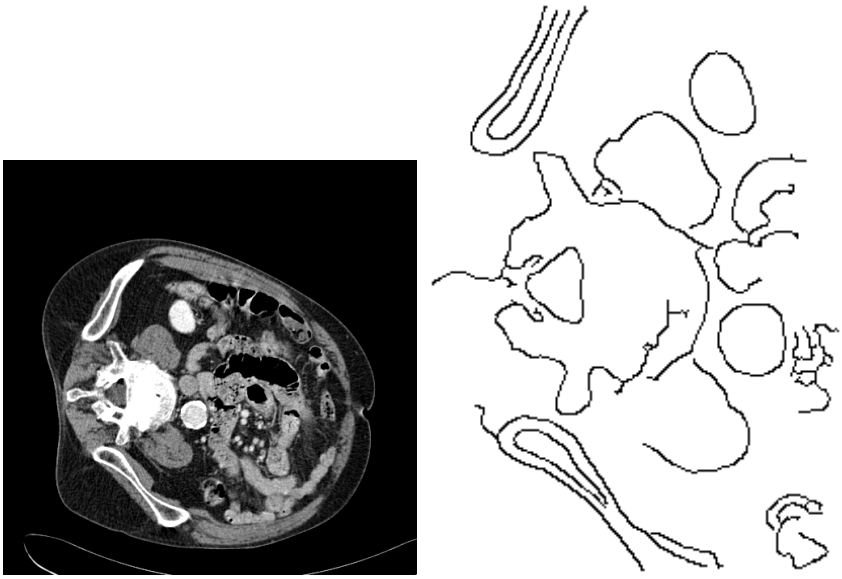
(a) US-image. (b)

(c) (d)

Figure A.4: (a) The US-image. (b) The edges detected by using P_{sc}. (c) The edges detected by using P_{sp}. (d) The edges detected by using P. For all the representation a natural threshold has been used.

A.5 CT-image.

The image in Fig.(A.5a) is a CT image. The representation in Fig.(A.5b) is a segment of the CT-image in the proximity of the spine. In order to detect the objects in the segment it suffices to use P_{sp}. P_{sp} is suitable since the contour of the objects is sharp and the objects are of large strength.



(a) CT-image. (b)

Figure A.5: (a) The CT-image. (b) The edges detected by using P_{sp} at $s = 2$.

Appendix B

The Fourier-Transform and Convolution.

Definition B.0.1 (1-D Fourier-Transform. [1].) *For a function $f \in L^1(\mathbb{R})$ the Fourier-transform is;*

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t}\, dt \tag{B.0.1}$$

□

Proposisjon B.0.1 (Derivation. [1],[9].) *If $f \in C^n(\mathbb{R}) \cap L^1(\mathbb{R})$ and if all the derivatives $f^{(k)}$ for $k = 1, 2, ..., n$ are in $L^1(\mathbf{R})$, then*

$$\hat{f}^{(k)}(\omega) = (i\omega)^k \hat{f}(\omega), \qquad \text{for } k = 1, ..., n. \tag{B.0.2}$$

□

Proposisjon B.0.2 (Convolution and Differentiation, [9].) *Let f be in $L^1(\mathbb{R})$ and let g be in $C^p(\mathbb{R})$. Assume $g^{(k)}$ is bounded for $k = 0, ..., p$. Then $f * g \in C^p(\mathbb{R})$ and $(f * g)^{(k)} = f * g^{(k)}$ for $k = 1, ..., p$.*

□

Definition B.0.2 (2-D Fourier-Transform. [1]) *For a function $f \in L^1(\mathbb{R}^2)$ the 2-D Fourier-transform is given by;*

$$\hat{f}(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i\omega_x x} e^{-i\omega_y y}\, dxdy \tag{B.0.3}$$

□

Proposisjon B.0.3 (Separable. [1]) *If $f \in L^1(\mathbb{R}^2)$ and there exists $h, g \in L^1(\mathbb{R})$ such that $f(x,y) = h(x)g(y)$, then $\hat{f}(\omega_x, \omega_y) = \hat{h}(\omega_x)\hat{g}(\omega_y)$.*

□

Proposisjon B.0.4 (Scaling. [1]) *If $f \in L^1(\mathbb{R}^2)$, and assume $s > 0$. Then the Fourier-transform of $f(t/s)$ equals $s\hat{f}(s\omega)$.*

□

Proposisjon B.0.5 (Gaussian. [1]) *If $f(t) = e^{-at^2}$, then $f(\omega) = \sqrt{\frac{\pi}{a}} e^{-a\omega^2/4a^2}$.*

□

Proposisjon B.0.6 (Convolution and 2-D Fourier-Transform. [1]) *If $f \in L^1(\mathbb{R}^2)$ and $h \in L^1(\mathbb{R}^2)$. Let $g(x,y) = f * h(x,y)$. Then $\hat{g}(\omega_x, \omega_y) = \hat{f}(\omega_x, \omega_y)\hat{h}(\omega_x, \omega_y)$.*

□

Proposisjon B.0.7 (N-dim Convolution and Differentiation, [10].) *If $f \in L^1(\mathbb{R}^n)$, $g \in C^k(\mathbb{R}^n)$, and $\partial^\alpha g$ is bounded for $\lfloor \alpha \rfloor \leq k$, then $f * g \in C^k(\mathbb{R}^n)$.*

□

Proposisjon B.0.8 (Boundedness by convolution,[10]) *If $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$, then $f * g$ exists a.e. $f * g \in L^p(\mathbb{R}^n)$ and $\| f * g \|_p \leq \| f \|_1 \| g \|_p$.*

Bibliography

[1] Mallat Stephane *a Wavelet tour of signal processing. Second Edition* Academic Press, USA. 1999

[2] Ward Cheny, David Kincaid *Numerical Mathematics and Computing. Fifth Edition* Brooks/Cole Pub Co, 2004

[3] Canny John *A Computational Approach to Edge Detection.* IEEE Trans. Patt. Anal. and Mach. Intell., 36:961-1005, September 1986

[4] Mallat Stephane, Hwang Wen Liang *Singularity Detection and Processing with Wavelets.* IEEE Trans. Info. Theory, 38(2):617-643, March 1992.

[5] Mallat Stephane, Zhong Sifen *Characterization of Signals from Multiscale Edges.* IEEE Trans. Patt. Anal and Mach. Intell., 14(7):710-732, July.

[6] Mallat Stephane, Hwang Wen Liang *Singularities and Noice Discrimination with Wavelets.* Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on. Vol. 4, 23-26 March 1992, 377-380

[7] Lu Jian, Weaver John, Healy Dennis, Xu Yansun *Noice Reduction with a Multiscale Edge Representation and Perceptual Criteria.* Time-Frequency and Time-Scale Analysis, 1992, Proc. of the IEEE-SP International Symposium, 4-6. Oct. 1992, 555-558

[8] Marr D, Hildreth E *Theory of edge detection.* Proc. Royal Soc. London, vol 207, pp. 187-217. 1980

[9] Gasquet C, Witomski P *Fourier Analysis and Applictions.* Springer-Verlag, 1999

[10] Folland Gerald B. *Real Analysis, Modern Techniques and Their Applications. 2.nd edition* Wiley Inter-Science 1999