DOI:10.15774/PPKE.ITK.2020.007

Ultrasound Image Modelling and Resolution Enhancement

Ákos Makra

Supervisor: Dr. Miklós Gyöngy

Pázmány Péter Catholic University

Faculty of Information Technology and Bionics

Tamás Roska Doctoral School of Sciences and Technology

A thesis submitted for the degree of *Doctor of Philosophy*

2020

DOI:10.15774/PPKE.ITK.2020.007

Contents

DOI:10.15774/PPKE.ITK.2020.007

DOI:10.15774/PPKE.ITK.2020.007

DOI:10.15774/PPKE.ITK.2020.007

List of Figures

DOI:10.15774/PPKE.ITK.2020.007

DOI:10.15774/PPKE.ITK.2020.007

DOI:10.15774/PPKE.ITK.2020.007

List of Tables

DOI:10.15774/PPKE.ITK.2020.007

Glossary

Notation	Description	Page List
.	element-wise multiplication	29
$\otimes$	correlation operator	33
∇	nabla operator	19
∇^2	the Laplacian, $\nabla^2 = \nabla \cdot \nabla$	22
$\Pi(.)$	rectangular function	39
$\text{III}(.)$	Dirac comb (Shah function)	39
$\mathcal{F}\{.\}$	Fourier transform	29
$\mathcal{F}^{-1}\{.\}$	inverse Fourier transform	47
$\|.\|$	norm	34
$\measuredangle$	phase of the complex functions	32
$\Delta\kappa$	deviation from mean compressibility at equilibrium	23
$\Delta\rho$	deviation from mean density at equilibrium	23
Λ	appropriate weight for a desired sparsity	34
β	bulk modulus, inverse of compressibility κ	21
γ	p-value for the l_p norm, $0 \leq p \leq 2$	34
κ_t	total compressibility	23
κ_v	equilibrium compressibility	23
κ	compressibility	21
λ	wavelength	8
$\phi(\mathbf{r}, t)$	scattered field	24
$\psi(.)$	Lucy–Richardson function	33
ρ_0	mean density at equilibrium	18
ρ_t	total density	18
ρ_v	equilibrium density	23
ρ	density	26
σ	variance of the noise	33

DOI:10.15774/PPKE.ITK.2020.007

Notation	Description	Page List
τ	spatial pulse length	12
σ_u^2	variance of all the pixel values of image u	48
a	scaling term	48
AC	autocorrelation	44
ADMM	alternating direction method of multipliers	34
ANN	artificial neural network	35
ARMA	autoregressive moving average	31
b	bias	35
B_{real}	real B-mode US image	48
B_{sim}	simulated B-mode US image	48
BD	beam diameter	10
BW	bandwidth	39
c	speed of sound	8
$\text{cycles}_{\#}$	number of cycles in the pulse	12
CNN	convolutional neural network	36
dx	interpixel-distance of macro photograph	52
D	diameter of the aperture	8
DL	deep learning	35
e_i	input voltage	27
E	additive error term	48
f	frequency	8
$f(t)$	total pressure acting on the transducer	27
$f(z)$	activation function	35
$f_{\#}$	f-number	11
F_L	focal length	11
FWHM	full-width at half maximum	10

DOI:10.15774/PPKE.ITK.2020.007

Notation	Description	Page List
g	optimal filter	32
G	Fourier transform of g	32
h	transfer function (PSF)	29
h_{mp}	1-D minimum-phase estimate of h	30
h_{pe}	geometric response of the ultrasound transducer	41
h_s	pressure impulse response of the scatterer	27
h_t	velocity potential impulse response	27
H	Fourier transform of transfer function h	29
H^*	complex conjugate of H	33
$i\,(z)$	backscattered signal in the axial direction	38
$I\,(u)$	Fourier transform of $i\,(z)$	39
I_{real}	real US image (post-beamformed RF)	48
I_{sim}	simulated US image	48
J	cost function	34
ML	machine learning	35
MSE	mean squared error	32
n	noise	29
$\tilde{\mathbf{n}}$	normal unit vector	19
N	near field distance	8
	Fourier transform of the noise n	29
NRMSE	normalized root mean square error	80
NSR	noise-to-signal ratio, inverse of SNR	33
p	(acoustic) pressure	17
$p\,(z)$	PSF in the axial direction	38
p_0	mean pressure at equilibrium	18
p_i	incident presssure	25

DOI:10.15774/PPKE.ITK.2020.007

Notation	Description	Page List
p_r	pressure on the transducer surface	27
p_s	scattered presssure	25
p_t	total pressure	18
$P(u)$	Fourier transform of $p(z)$	39
PSF	point-spread function	28
$\mathrm{PSF}_{\mathrm{data}}$	data-based PSF	53
$\mathrm{PSF}_{\mathrm{deconvolution}}$	deconvolution-based PSF	53
$\mathrm{PSF}_{\mathrm{est}}$	estimated PSF	48
$\mathrm{PSF}_{\mathrm{FieldII}}$	Field II-based PSF	53
PSNR	peak signal-to-noise ratio (peak SNR)	81
$\mathbf{r}$	position vector	17
$\mathbf{r}'$	location of acoustical disturbance	25
$\mathbf{r}_0$	location of single scatterer	26
$\mathbf{r}_r$	location of transducer surface	27
R	reflection coefficient	26
$R(f)$	RAMP function (function of frequency)	67
R^2	coefficient of determination	46
R_A	axial resolution	11
R_B^2	R^2 between the simulated and real B-mode US images	48
R_I^2	R^2 between the simulated and real US images	48
R_L	lateral resolution	11
RAMP	frequency-weighted axial filtering	67
RF	radio-frequency	14
s	distance	14
$s(z)$	SF in the axial direction	38
S	surface area	19

DOI:10.15774/PPKE.ITK.2020.007

Notation	Description	Page List
$S(u)$	Fourier transform of $s(z)$	39
S_{kernel}	size of the kernel	38
$S_{lim}(u)$	band-limited $S_{lim}(u)$	39
S_N	power spectrum density of the noise	33
S_r	area of the receiver the transducer	27
S_t	area of the transmit the transducer	27
S_X	power spectrum density of the signal	33
SAM	scanning acoustic microscopy	15
SBE	step-based estimation	77
SF	scattering/scatterer function	2
$\mathrm{SF}_{\mathrm{est}}$	estimated SF	48
$\mathrm{SF}_{\mathrm{points}}$	point-based SF from $\mathrm{SF}_{\mathrm{project}}$	51
$\mathrm{SF}_{\mathrm{project}}$	projection-based SF from $\mathrm{SF}_{\mathrm{threshold}}$	51
$\mathrm{SF}_{\mathrm{threshold}}$	threshold-based SF	51
SNR	signal-to-noise ratio, inverse of NSR	42
SR	super-resolution	3
t	time	17
t_p	elapsed time between the emitted pulse and detected echo	14
TV	total variation	34
u	angular frequency	39
u_0	central frequency	39
US	ultrasound	1
$\mathbf{v}$	particle velocity	17
v_{pe}	electric response of the ultrasound transducer	41
V	volume	19
w_0	weight parameter	35

DOI:10.15774/PPKE.ITK.2020.007

Notation	Description	Page List
w_r^δ	electromechanical response of the receive transducer	27
w_t^δ	electromechanical response of the transmit transducer	27
x	lateral distance	5
	underlying structure (SF)	29
$\hat{x}$	estimated input	32
x_0	neuron input	35
$\hat{x}_k$	the estimate of x after k number of iterations	33
X	Fourier transform of underlying structure x	29
$\hat{X}$	Fourier transform of estimated input $\hat{x}$	32
y	elevation distance	5
	observed image	29
y_{max}	maximum value of image y	81
y_{min}	minimum value of image y	81
Y	Fourier transform of observed image y	29
z	axial distance	6
Z	acoustic impedance	26
Z_1	acoustic impedance of the first medium	26
Z_2	acoustic impedance of the second medium	26
ZP	zero-phase (estimate)	43

DOI:10.15774/PPKE.ITK.2020.007

0.1 Declaration of original work used

This section introduces the main references of the current thesis. In Chapter 1, Section 1.1.2 is based on [Au1], Section 1.1.3 is based on [Au2, Au3]. In Chapter 2, Section 2.1 is based on the Master's Thesis of the author [Au4], Section 2.2 is based primarily on the Bachelor's Thesis of the author [Au1], in which Section 2.2.7 is based on [Th1]. Furthermore, Sections 2.3.1, 2.3.2 and 2.3.4 are based on [1] [Au2], and Section 2.3.6 is based on [Au5]. Chapter 3 is based on [Th1], Chapter 4 is based on [Th2], and Chapter 5 is based on [Th3].

Further references, which are mostly independent of the author, are provided in the appropriate (sub)sections. To facilitate differentiation, pictures taken verbatim from the publications of the author are indicated by a standalone citation at the end of the figure caption, whereas modified/adapted or images taken verbatim from other sources are annotated in a normal manner.

DOI:10.15774/PPKE.ITK.2020.007

Chapter 1

Introduction

1.1 Motivation

1.1.1 Relevance of ultrasound imaging

Diagnostic ultrasound has been in use for 60 years now and it has become one of the most popular medical imaging methods nowadays. Diagnostic ultrasound imaging commonly utilizes frequencies in the range of 3–20 MHz. The use of higher frequencies limits the depth of penetration, however it also increases resolution.

Recently, ultrasound (US) has been actively used not only for medical diagnostic purposes [2–4], but also for high-intensity focal beam surgery to produce precise and selective damage to tissues [5–7], biometric recognition [8], non-destructive testing [9–19], and has many applications in the food industry [20–23] among others. Its wide range of applications stems from its numerous advantages such as cost-effectiveness, portability, and using non-ionizing radiation compared to many other procedures such as X-ray, CT or PET, all of which are using potentially harmful radiation. On the other hand, the interpretation of US images is still quite a subjective task despite the numerous quantitative US studies [24–33].

The connection between the fine microscopic structure of tissues and the resulting US image is at present not fully understood, which further motivates the development and the importance of validating image formation models.

DOI:10.15774/PPKE.ITK.2020.007

1.1.2 Relevance of validation of image formation models

Interpretation of US images is quite a subjective task, so the mapping of the various histological pathologies is an empirical (and implicit) procedure for the radiologist experts. There are many theoretical models to describe US imaging, however, there is little research to validate these models. The following paragraphs attempt to categorize these approaches.

First, for US imaging to occur, US needs to propagate to the scatterers in question, then scattering of the incident wave must occur, and the scattered wave must propagate back to the transducer. The difference lies in how these phenomena are treated in US imaging models.

Propagation is usually considered to be linear, which is based on the assumption that the deviations in pressure and density that support the propagation of the wave (for more details see Section 2.2.4) are small relative to the mean pressure and density. If the medium is homogeneous, the waves travel through unimpeded; however, any degree of inhomogeneity causes ultrasound scattering, which arises from perturbations in density and compressibility [34]. Considering the backscatter of ultrasound in the direction of the original incident wave (180° scattering), the scattering function (SF) in terms of density and compressibility may be reduced to a SF (or alternatively, tissue reflectivity function) expressing relative changes in acoustic impedance, which is consistent with the 1-D model of wave reflection [35, 36] [37, pp. 304–306]. Other authors have opted to express this scattering function in terms of changes in the bulk modulus [38].

Another issue concerning modeling is the nature of the scattering. Most models neglect multiple scattering and assume that the scattered field is generated only by the incident field, an assumption known as the Born approximation (see Section 2.2.5). However, there is a conceptual split in research that treats the scattering medium as consisting of discrete scatterers [34, 38] and those that regard it as a continuously varying acoustic maps [39].

Another simplification is to assume the impulse response of the scatterers spatially invariant [40], which means that the position of the scatterer relative to the US transducer is irrelevant in the terms of impulse of the scattering. This assumption

DOI:10.15774/PPKE.ITK.2020.007

is called shift-invariance. For more detail about the validity of this simplification, the reader is directed to [41] and Chapter 4.

If as a first step, it would be possible to validate an image formation model with simple, inanimate scatterers, it would open the way toward exploring the relationship between histology and US images.

1.1.3 Relevance of resolution enhancement

Imaging modalities of any kind have a theoretical limit on their feasible resolution. The objective of the super-resolution (SR) algorithms is to break this boundary, thereby obtaining an image of higher quality with the same physical setup.

There has always been a great demand for producing images with better and better resolution, either by creating a better physical setup, or using post-processing techniques, whether it is about security cameras [42–44], satellites [45–50], professional photography [42, 51–53] or even the HUBBLE space telescope [54–57]. The same rules apply for medical purposes: the higher the resolution of an image, the more precise the diagnosis.

Concerning software-based methods for enhancing image resolution, the algorithm can be used either on sub-pixel-shifted frames by stacking them, or as a post-processing step where even one frame can be satisfactory. The use of SR techniques provides the possibility of receiving a more detailed image at a lower cost compared to the expensive and time-consuming process of building a new hardware capable of delivering the same quality.

Nevertheless, along with other imaging modalities (such as MR, CT or light microscopy) its resolution is heavily dependent on the wavelength (higher frequency, thus shorter wavelength leads to better resolution), which in the case of sound is a lot poorer than that of light or X-ray. The transducer and its frequency also determine the penetration depth (the higher the frequency, the smaller the mentioned depth is) [58, p. 116]. To be able to examine deeper layers of the medium, lower frequencies should be used, which, however, decreases the resolution.

Taking into account the benefits of US imaging it would be worthwhile if the image resolution and signal-to-noise quality could be improved by post-processing

DOI:10.15774/PPKE.ITK.2020.007

methods. This work is concentrated around US images; however, the algorithms to be presented can be adapted to other imaging modalities as well.

1.2 Overview of current thesis

In this section a brief overview of the current thesis is given.

In Chapter 2, the theoretical basis of ultrasound image formation and resolution enhancement is described. In Section 2.2, the theoretical background of acoustics and the basic concepts regarding US are introduced, including the shift-invariant convolution model. In Section 2.3, the mathematical background of resolution enhancement is provided. Further in this section, the classical deconvolution techniques used for image resolution enhancement are described in Section 2.3.3, with cost function minimization in Section 2.3.4. The theoretical background of deep learning is introduced in Section 2.3.5. Last, the current US specific resolution enhancement methods are discussed in Section 2.3.6, with an emphasis on equivalent scatterers and axial processing.

The next chapters introduce the scientific work of the author. In Chapter 3, a method to experimentally assess the accuracy of the previously described shift-invariant convolution model (see Section 2.2.7) is presented. In Chapter 4, the resolution enhancement of ultrasound B-mode images using deconvolution (see Section 2.3.3) and axial processing (see Section 2.3.6) is presented, whereas Chapter 5 demonstrates the resolution enhancement of ultrasound C-scan images using deep learning (see Section 2.3.5).

In Chapter 6, a summary is given about the new scientific results of the author in the form of thesis points.

DOI:10.15774/PPKE.ITK.2020.007

Chapter 2

Theory

2.1 Basics of ultrasound imaging

The aim of the current section is to provide the reader with the basics of ultrasound imaging and introduce the main concepts regarding the topics covered in this thesis.

Mechanical waves with frequencies higher than 20 kHz are called ultrasound. Its propagation results in a periodic change of pressure in space and time. Ultrasonic diagnostics take advantage mainly of the fact that ultrasound waves are reflected from different interfaces in the medium: the elapsed time and the intensity of the reflection can provide a good estimate about the physical properties of the examined medium.

As with any imaging modality, there are a set of different parameters which together adequately describe the system. In general, these parameters are not independent of each other, and the choice of parameters greatly determines the application area. A simple example is when the central frequency of the transducer is increased, which (assuming an increase in bandwidth) leads to greater resolution but at the same time decreases penetration depth and hence more superficial examination of the tissue. Given the above consideration, it is crucial to have a clear understanding of the basics of image formation and to be aware of the connections and trade-offs between the system parameters.

Conventionally, the lateral, elevation, and axial directions are denoted by x, y,

DOI:10.15774/PPKE.ITK.2020.007

and z, respectively, unless otherwise noted.

The following sections (Sections 2.1.1 to 2.1.4) are primarily based on the following sources [37, 58, 59] unless otherwise annotated.

2.1.1 Transducer

A transducer is a device capable of transforming energy from one form to an another. In the case of US, it converts the electrical energy into mechanical pressure waves and vica versa, by utilizing the inverse and direct piezoelectric effects (discovered by the Curie brothers in 1880 [60] [58, p. 3]), respectively. It has two different modes, namely transmit and receive. During transmission mode it emits (a series of) sound waves into the body by converting electrical energy to pressure vibrations (inverse piezoelectric effect). In receive mode it captures the backscattered sound waves and interprets those as electrical signals (direct piezoelectric effect).

The fundamental part of the transducer is the piezoelectric element itself. In general, it is made of a crystal-like material like tourmaline, quartz, topaz or cane sugar [61]; however, synthetic materials, like poly-vinylidene fluoride [62] are capable of producing the piezoelectric effect as well.

Since the discovery of the piezoelectric effect many forms of the transducer have been invented, each of them suitable for different tasks. In this section the main types of transducers will be introduced.

Transducer types

1. Single-element transducers: The simplest form of transducers are made of a single element. For focusing, two methods may be used. One of them focuses the beam using an acoustic lens, which for instance can be made of plastic, epoxy, liquid or rubber [63]. The other method is to press the element itself into a curved shape [64]. Since single-element transducers only have a fixed focus, in order to capture an image manual scanning has to be performed. Note that even a so-called unfocussed (flat-surface) single-element transducer has a natural focus, which is defined by Eq. (2.1). In addition, the lateral resolution

DOI:10.15774/PPKE.ITK.2020.007

changes with depth, which can lead to a false impression of inhomogeneity even in the case of a homogeneous sample.

2. Linear array transducers: The second main form of transducers is the linear array. It contains a line of elements (for instance 320 pieces, spaced over an aperture of 56 mm [65]). The smaller the face of the transducer, the more divergent the beam will become during propagation. In order to decrease this divergence, a subset of the elements (8-16) is selected and pulsed simultaneously. For the next line the neighbouring subgroup will be activated (shifting the previous elements by one) and so on. However, due to the shifting sub-aperture, the physical footprint of this arrangement is sometimes too large to scan through relatively small acoustic windows (such as in imaging the heart through the ribs).

3. Phased array transducers: In the phased array, the elements are again placed in a row like in the case of the linear array. However, this time, all the elements are excited together but with small time-delays, needing independent delay circuits. This manner of emission of sound waves results in a curved propagating wavefront. By changing the length of delays, the beam can be steered into various directions and with different focal lengths. When receiving the signal, the same delay factors can be used in order to have dynamic focusing, meaning all points along the z-axis will be in focus. Since there is no shifting sub-aperture, the footprint of the transducer can be as small as 2 cm. More generally, if the elements are arranged on a planar grid, the beam can be steered in three dimensions, resulting in 3D ultrasonic imaging. By combining linear and phased arrays, the reduction of beam-divergence and focusing can be accomplished at the same time.

4. Annular array transducers: The last common transducer type to be discussed in this introduction is the annular array. In this case the elements are arranged as a set of concentric rings, possibly evenly placing them at different heights, giving a spherical curvature to the array. However, this curvature can be also achieved by activating the circle-elements with time-delays. It also records

DOI:10.15774/PPKE.ITK.2020.007

a single line at a time as single-element transducers, but gives a better and possibly changeable focus, and greater depth of field as well. Effectively, a large number of fixed-curvature spherical transducers can be placed farther away, thus they can be used with higher frequency. The limit which needs to be considered when placing the elements is the effect of grating lobes, which can be overcome by creating the transducer elements to be smaller than half of the used wavelength. In phased arrays, the creation of such small electrical circuits is always a challenge; however, annular arrays do not suffer from this restriction.

2.1.2 Beam parameters of spherically focussed transducers

Spatial distribution of the ultrasound field

The spatial distribution of the US field (otherwise referred to as the ultrasound beam) is generated as a result of the inverse piezoelectric effect, when applied voltage is transformed into a periodically varying pressure (sound) wave. The beam can be divided into two regions: near field (Fresnel zone) and far field (Fraunhoffer zone). Figure 2.1 shows that in the near field the pressure intensity reaches multiple local minima and maxima and ends in a last, global maximum, called near field distance. The distance at which the far field begins can be calculated by the following formula [59, p. 43]:

$$N = \frac{D^2 f}{4c} = \frac{D^2}{4\lambda}\,, \tag{2.1}$$

where N means the near field distance, D is the diameter of the aperture, f is the frequency, c stands for the propagating speed of sound in the medium, and λ annotates the wavelength. It is important to note that the near field is the same for both unfocussed and geometrically focussed transducers. However, the focal length will differ, meaning that the focal length will coincide with the near field distance for an unfocussed transducer.

In the far field the beam becomes divergent, causing the lateral resolution to heavily decrease. Geometrically focussed transducers with help of an acoustic lens will cause the beam width to decrease at a certain distance (focal length) from the

DOI:10.15774/PPKE.ITK.2020.007

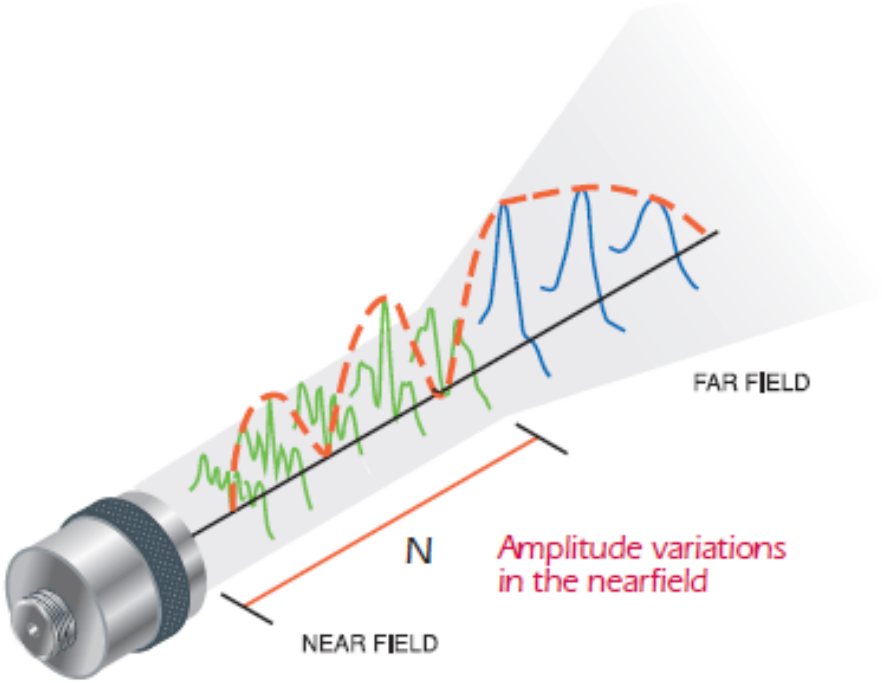

Figure 2.1: Sound field of a single-element transducer. The emitted beam can be divided into two parts: the near field (Fresnel zone) and the far field (Fraunhoffer zone). The near field starts directly after the transducer surface, where the amplitude of the pressure wave reaches several local minima and maxima, and ends in a last, global maximum. The location of the last point defines the near field distance (which in the case of geometrically unfocussed transducers is the same as the focal length), and is annotated by N in the figure. Beyond the near field distance begins the far field, where the intensity of the sound wave decreases progressively to zero due to the divergent nature of the beam. Picture taken from [59, p. 43].

transducer, which distance, as a rule of thumb, is approximately half of the radius of the curvature. As a result, at the focal length the lateral resolution will improve, as well as the extent of divergence of the beam in the far field, but the depth of field will be decreased.

In addition to the relatively simple concept of a transmitted beam, an analogous concept exists for the so-called receive beam, which is the spatial distribution of the receive sensitivity of the ultrasound transducer. Where the same transducer is used for transmit and receive, the transmit and the receive beams can generally be considered to have the same shape.

DOI:10.15774/PPKE.ITK.2020.007

2.1.3 Axial and lateral resolution

The beam diameter (BD), which is equivalent and often referred to as beam width, plays an important role when it comes to the sensitivity of a transducer. The beam consists of side lobes and a main lobe, which together are considered as the beam. Side lobes are essentially off-axis energies which can lead to different artefacts and are dependent on many factors such as the frequency, element spacing, number of elements and bandwidth. At the point of interest the smaller this diameter is the more energy is concentrated onto that point, which means the more energy will be reflected, resulting in a higher lateral resolution.

To define beam width, it is needed to define the full-width at half maximum (FWHM) [37, p. 173]. Considering a distribution function (such as pressure amplitude $\Delta p(x)$), as it can be seen in Fig. 2.2, around its maximum value such an interval can be found along the x-axis, where the value of the function is higher or equal than the half of the maximum value. In the literature, generally the -6 dB pulse-echo criteria is used, which is an equivalent form due to the fact that it describes such a relationship where the voltage amplitude of the output is half of the input. This interval along the x-axis defines the so called FWHM.

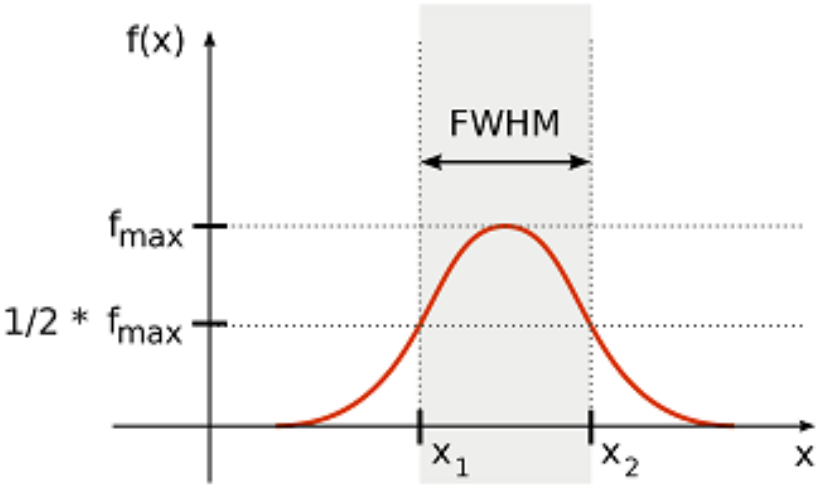

Figure 2.2: Full-width at half maximum. Considering a distribution function, the so-called full-width at half maximum means such an interval along the x-axis, where the value of the function is equal or higher than the half of the maximum. Note that $f(x)$ represents a voltage-like function, such as pressure amplitude $\Delta p(x)$. Picture taken from [66].

DOI:10.15774/PPKE.ITK.2020.007

The BD varies along the main lobe axis, as it is typically wide (in a perpendicular plane to the propagation direction) in the near and far fields and is the smallest at the focal length. The -6 dB pulse-echo BD at the focal length can be estimated, taking into account the FWHM definition and the Sparrow criterion [67, p. 189], as follows [59, p. 43]:

$$BD = 1.02\lambda \cdot \frac{F_L}{D} = 1.02\lambda \cdot f_{\#}\,, \tag{2.2}$$

where F_L stands for the focal length, D is the diameter of the aperture, and $f_{\#}$ is the f-number of the transducer.

On one hand, as it has been described above, the BD will determine the lateral resolution of the system (the lateral size of the received signal), as shown in Fig. 2.3 a). On the other hand, this resolution describes what is the smallest distance between two scatterers – in a perpendicular plane to the propagation – which can still be distinguished (see Fig. 2.3 b)). It essentially means when two scatterers are in the same plane and both of them are inside the main lobe, they cannot be distinguished; however, if one of them is located outside, the system is capable of identifying them as two different structures.

The lateral resolution R_L can also be derived from Rayleigh scattering, when a scatterer is much smaller compared to the wavelength. In this case the lateral resolution can be estimated (taking into account the Rayleigh criterion [67, p. 187]) as follows [37, p. 173]:

$$R_L = 1.22\lambda \cdot \frac{F_L}{D} = 1.22\lambda \cdot f_{\#}\,, \tag{2.3}$$

where the multiplication factor of 1.22 is present due to the Bessel functions [37, p. 173] and Rayleigh criterion [67, p. 187].

The axial resolution R_A is also an important parameter of an imaging system. It describes what is the smallest distance between two scatterers located parallel to the beam, which can still be distinguished (see Fig. 2.4). It is determined by the spatial pulse length, which is the number of cycles in the pulse multiplied by the wavelength of the propagating pressure wave, and can be estimated as follows [37, p. 511]:

$$R_A = \frac{1}{2}\lambda \cdot \text{cycles}_{\#} = \frac{1}{2}\tau\,, \tag{2.4}$$

DOI:10.15774/PPKE.ITK.2020.007

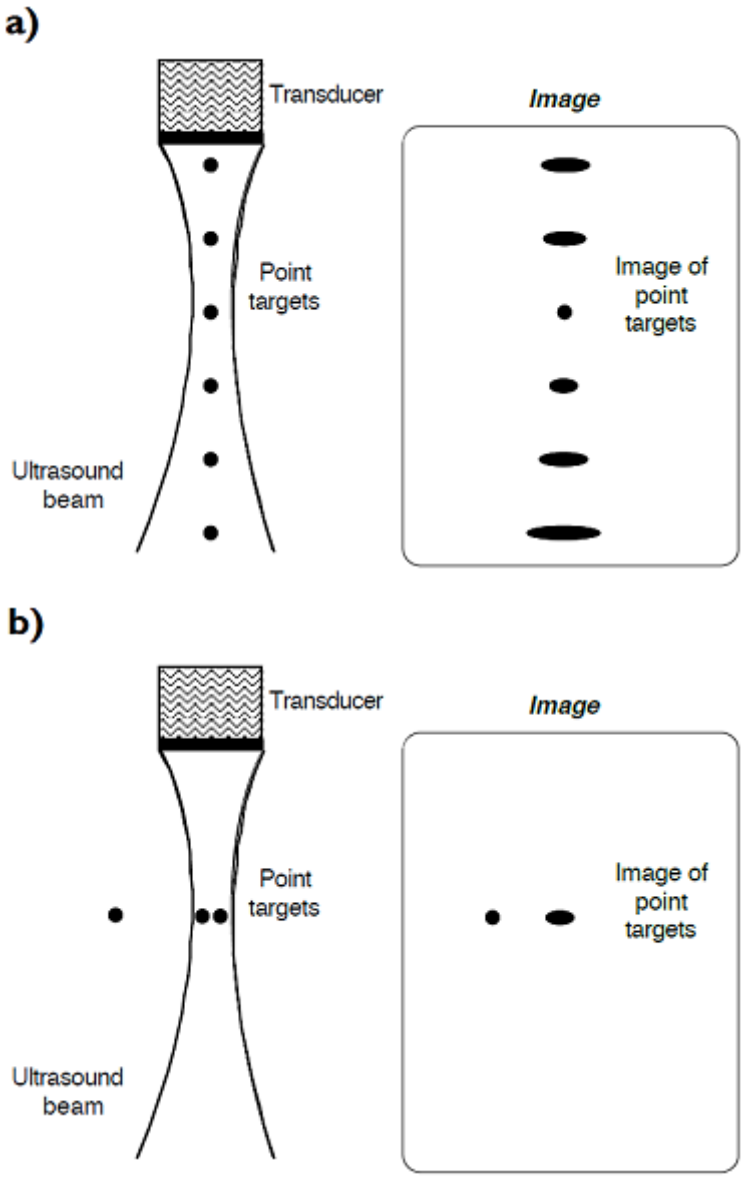

Figure 2.3: Lateral resolution. Figure 2.3 a) shows that how scatterers at different locations on the main axis would be present in an ultrasound image with the given beam profile. The farther the scatterers are from the focal length where the BD is the smallest, the laterally distorted they appear in the actual ultrasound image. In addition, in Fig. 2.3 b) it can be seen when two scatterers are inside the beam in the focal length plane, and in that case they cannot be distinguished from each other as two different scatterers. On the other hand, if one of those scatterers is outside the beam, they will appear in the ultrasound image as different objects. Picture adapted from [68, p. 13].

where λ is the spatial wavelength, $\text{cycles}_{\#}$ is the number of cycles in the pulse, and τ represents the spatial pulse length.

DOI:10.15774/PPKE.ITK.2020.007

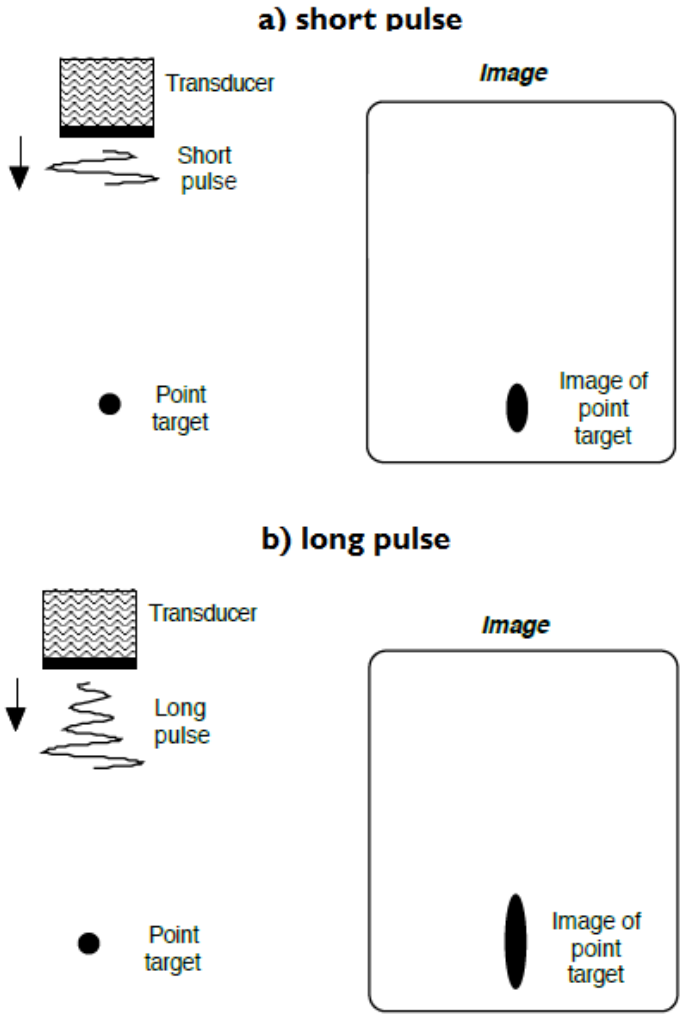

Figure 2.4: Axial resolution. The axial resolution is determined mainly by the length of the excitation pulse (or the so-called number of cycles parameter). In Figure a) it can be seen that the axial resolution is decent in case of using a short pulse as excitation. Figure b) shows a big amount of axial distortion in case of using a long pulse. Picture adapted from [68, p. 14].

2.1.4 Ultrasound imaging modes

There are different ultrasound imaging modes such as A-mode, B-mode, B-flow, Doppler, Power Doppler, C-mode, or M-mode among others; however, this work is based on the conventional A-mode, B-mode and C-mode imaging. This section describes these methods.

DOI:10.15774/PPKE.ITK.2020.007

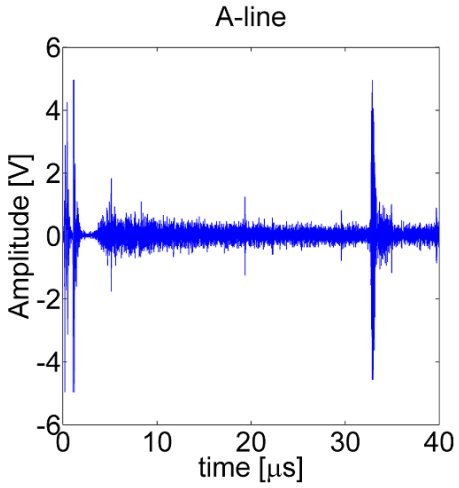

Figure 2.5: A conventional A-line. The first 5 μs shows an artefact of the imaging system due to the constant high frequency switch between transmit and receive modes. Around 32 μs a reflection can be observed from the boundary of two media having different physical properties. Picture taken from [Au4].

A-mode

The so-called A-mode means Amplitude-mode and is the simplest of all modes. It corresponds to the conventional oscilloscope display, when an emitted pulse is returned as an echo due to local inhomogeneities, and it is displayed as a function in a time-amplitude (or distance-amplitude) plane, where a higher amplitude means a greater acoustic impedance difference between the boundaries. Such an A-line can be seen in Fig. 2.5. Other sources refer to A-lines after taking the envelope of the post-beamformed radio-frequency (RF) signal. The distance from the surface of the transducer can be calculated knowing the propagating speed of sound in the medium as follows:

$$s = \frac{c \cdot t_p}{2}\,, \tag{2.5}$$

where s is the distance, c is the speed of sound in the examined medium, and t_p stands for the elapsed time between the emitted pulse and detected echo.

B-mode

A B-mode image shows a cross-section plane of the medium, which is parallel to the propagating waves. The so-called B-mode means Brightness-mode and orig-

DOI:10.15774/PPKE.ITK.2020.007

inates from the conventional display of US images, where the higher amplitude of a pulse-echo signal is, the brighter is the image at those areas. An actual B-mode image can be generated using a plethora of different insonification strategies such as focussed beams or diverging-wave transmissions [69]. It is also possible to record equidistant A-lines with a single-element transducer, *e.g.*, by using a scanning acoustic microscope system (SAM) (see Fig. 2.6), and use the envelope of the amplitude for the visualization to produce a B-mode image. The A-lines can be put next to each other based not only on distance, but time or angle as well.

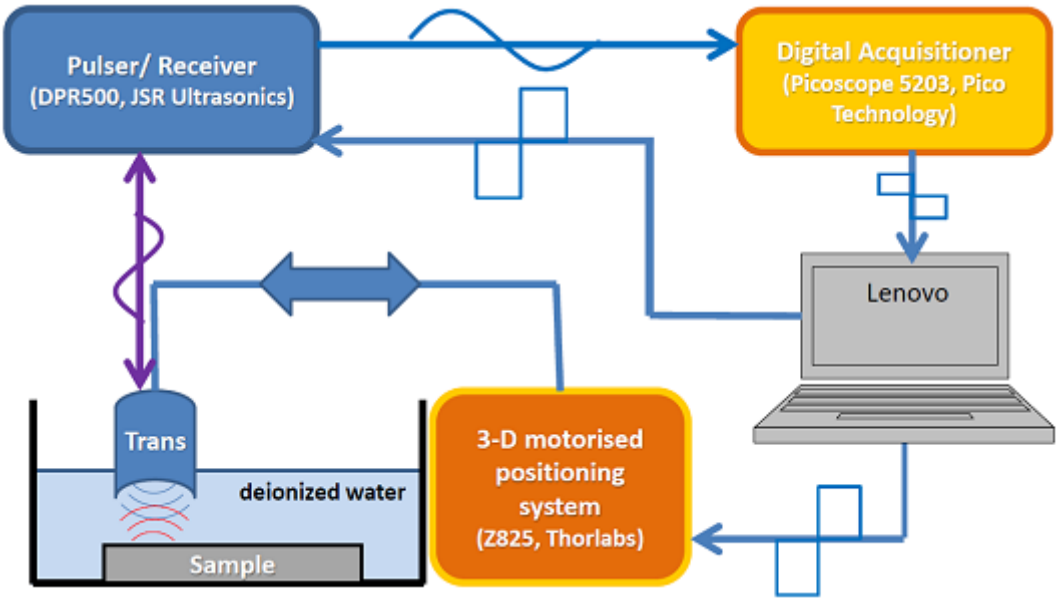

Figure 2.6: Schematic of a SAM system. During the movement of the transducer the impulse generator (or referred to as pulser) at a certain desired frequency receives a trigger signal, which generates a short electric impulse with a high amplitude. This electrical impulse will be transformed into a sound (pressure) wave, which is (generally) focussed on the surface of the sample. The scattered (reflected) wave will be detected, converted back to electric impulse and will get recorded by the digital acquisitioner. After saving the desired data, the transducer moves to the next position with the help of the micropositioner system. Picture taken from [Au4].

C-mode

C-mode imaging results in such a cross-sectional plane of the examined media, which is normal to the propagating waves (and therefore to the B-mode images). C-mode images are usually generated by a SAM system, where a motorized movement

DOI:10.15774/PPKE.ITK.2020.007

of a single-element transducer performs a scan over a pre-defined grid (usually in 2-D) and records an A-line at every grid point. Such a system can be seen in Fig. 2.6. By providing depth information over a 2-D plane the result is a 3-D volumetric data. Such A-lines can be either integrated over a given window axially or individual planes can be displayed selecting the same depth along every A-line.

One common form of SAM imaging involves taking the integral of the absolute values of A-lines in the axial direction, which is used for the visualization process, resulting in a C-scan image.

2.2 Theory of ultrasound image formation

2.2.1 Overview

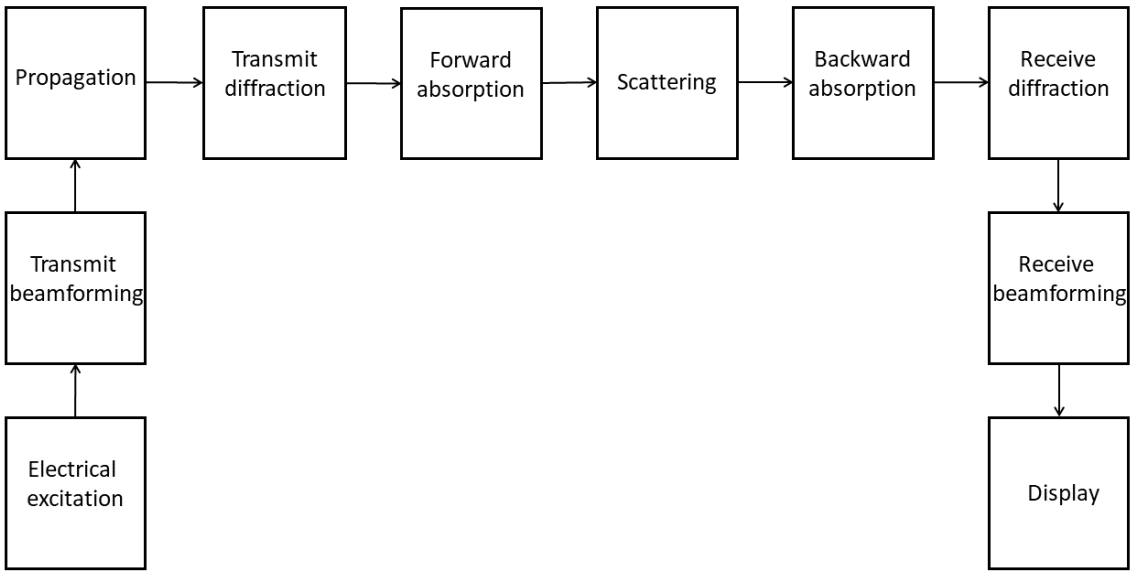

Figure 2.7: Block diagram of US image formation listing the major processes. Image is based on [70, p. 43]

For an ultrasound image to occur, there are many steps which have to be considered during the US image formation. Figure 2.7 lists the major processes. For more comprehensive system details the reader is directed to [70, pp. 43–45].

First, an electrical excitation is used on the transducer. The transmit beamformer by receiving this signal ensures that the appropriately delayed pulses arrive

DOI:10.15774/PPKE.ITK.2020.007

at the transducer elements. The array elements convert the electrical energy to pressure (sound) waves by utilizing the inverse piezoelectric effect (as mentioned in Section 2.1.1). The propagating waves are undergoing diffraction, absorption, and scattering (both travelling further away and back towards the transducer), during which its energy is decreased (this energy loss can be compensated for). Upon the waves reaching the surface of the transducer, the pressure is going to be converted back to electrical signals, taking into account the appropriate time delays. Before the display, the data can be post-processed with different techniques such as time gain compensation, speckle reduction, envelope detection or log-compression among others, to meet desired needs.

In the following, the physics behind the two most thesis-related processes will be introduced, namely the propagation (see Sections 2.2.2 and 2.2.3) and scattering (see Sections 2.2.4 and 2.2.5). Next, the whole image formation will be described as a shift-variant convolution-model, taking into account certain assumptions (see Section 2.2.6). Last, the shift-invariant convolution model will be introduced (see Section 2.2.7).

2.2.2 Governing equations of acoustics

In this section, the governing equations of acoustics, which are going to lead to the homogeneous acoustic linear wave equation, will be derived using the following sources: [37, Sections 1.3.5 and 5.4] [71, Section 2.1.1] [72, Chapter 2] [73, Chapter 1] [74, Section 3.5 and Chapter 16] [34].

Acoustics deals with describing the propagation of mechanical waves through different media. The acoustic wave equation is a second order partial differential equation, which establishes a relation between the acoustic pressure p and particle velocity $\mathbf{v}$ as a function of position $\mathbf{r}$ and time t.

To obtain a linearized form of the wave equation, it is crucial to use approximations, and further variables need to be defined. First of all, it is assumed that the acoustic pressure and density can be written as

$$p_t(\mathbf{r}, t) = p_0 + p(\mathbf{r}, t) \qquad (p \ll p_0) \tag{2.6}$$

DOI:10.15774/PPKE.ITK.2020.007

$$\rho_t(\mathbf{r}, t) = \rho_0 + \rho(\mathbf{r}, t) \qquad (\rho \ll \rho_0) \tag{2.7}$$

at any given location $\mathbf{r}$ and time t, in which p_t is the total pressure, p_0 is the mean pressure at equilibrium, p is the pressure fluctuation caused by the propagating wave, ρ_t is the total density, ρ_0 stands for the mean density at equilibrium and ρ is the density fluctuation induced by the wave. Figure 2.8 illustrates the situ-

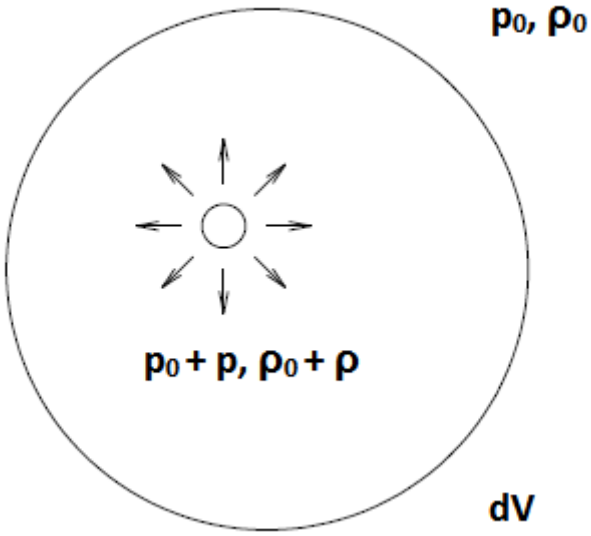

Figure 2.8: A volume of fluid dV with the local pressure p and local density ρ can be seen. The local pressure and density variations are generated by an acoustic disturbance (such as a propagating wave through the media). The mean pressure and density at equilibrium are denoted by p_0 and ρ_0, respectively. Picture adapted from [72, p. 6].

ation: there is a fluid volume dV with mean pressure p_0 and mean density ρ_0 at equilibrium. After any acoustic disturbance local pressure and density fluctuations can be observed, denoted by p and ρ, respectively. Considering these assumptions, the ultrasound waves cause perturbations around the mean density and pressure; furthermore, these perturbations are assumed to be small. To fully describe the acoustic field, the following three equations are needed to derive in linearized forms: equation of motion, continuity equation and equation of state.

DOI:10.15774/PPKE.ITK.2020.007

The equation of motion

The equation of motion is based on Newton's second law, forming a relationship between the force acting on the fluid particle's volume and its momentum's alteration. Considering a fluid volume V with pressure distribution p on its surface, the force vector acting on it can be described as

$$\mathbf{F} = -\iint_S p\ \tilde{\mathbf{n}}\ dS\,, \tag{2.8}$$

where S stands for the surface area which encloses the volume and $\tilde{\mathbf{n}}$ is the normal unit vector pointing outwards on the differential surface area dS. By using the Divergence Theorem[1] the surface integral can be transformed into a volume integral

$$\mathbf{F} = -\iiint_V \nabla p\ dV\,, \tag{2.9}$$

where ∇ is the well-known nabla operator. The equation states that the force per unit volume equals $-\nabla p$. According to Newton's second law it is known that the force per unit volume equals to mass per unit volume multiplied by acceleration (the derivative of velocity). Mass per unit volume is density, besides taking into account the assumptions of Eq. (2.7) the following form can be derived:

$$\rho\frac{\partial \mathbf{v}}{\partial t} \cong \rho_0\frac{\partial \mathbf{v}}{\partial t} = -\nabla p\,, \tag{2.10}$$

where $\mathbf{v}$ is the velocity vector of the fluid particle. This equation is called the linearized equation of motion.

The continuity equation

The continuity equation is about mass conservation and it is based on the fact that mass can not arise and can not disappear. Let us see Fig. 2.9. It can be described how many times more mass flows out per second than flows in over the surface area:

$$q_m = \iint_A \rho\mathbf{v}\mathbf{dA} \qquad \left[\frac{kg}{s}\right]. \tag{2.11}$$

[1]Divergence Theorem: also known as Gauss-Ostrogradsky Theorem

DOI:10.15774/PPKE.ITK.2020.007

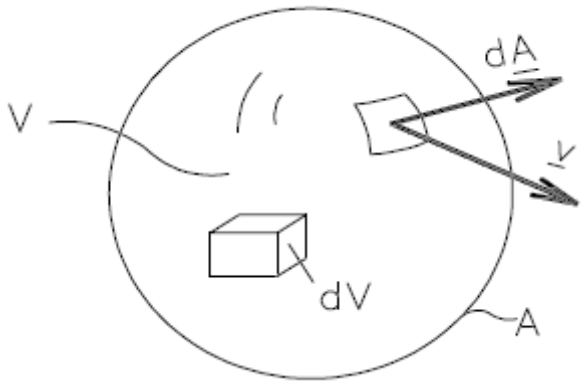

Figure 2.9: A volume of fluid V at a fixed location, enclosed by a surface A. Inside there is a differential volume dV, a normal vector of the surface $\mathbf{dA}$ *and the velocity vector of the outflowing mass* $\mathbf{v}$*. Picture taken from [74, p. 29].*

Outflow from the volume results in decreased mass; therefore, density is also decreased. The change of mass in the volume V can be determined by the following integral:

$$\iiint\limits_V \frac{\partial \rho}{\partial t} dV \,. \tag{2.12}$$

The normal vector of the surface element $\mathbf{dA}$ is pointing outwards; therefore, if Eq. (2.11) is positive, the mass is decreasing in the volume. Taking mass conservation also into account it can be derived that:

$$-\iiint\limits_V \frac{\partial \rho}{\partial t} dV = \iint\limits_A \rho \mathbf{v} \mathbf{dA} = \iiint\limits_V \nabla \left(\rho \mathbf{v}\right) dV \,. \tag{2.13}$$

The right-hand side of the equation was converted using the Divergence Theorem. Considering the integration is over the same volume and using obvious rearrangements, the following can be obtained:

$$\iiint\limits_V \left[\frac{\partial \rho}{\partial t} + \nabla \left(\rho \mathbf{v}\right)\right] dV = 0 \,. \tag{2.14}$$

The integral equals zero if and only if the integrand equals zero. Accordingly, it can be simplified to

$$\frac{\partial \rho}{\partial t} + \nabla \left(\rho \mathbf{v}\right) = 0 \,, \tag{2.15}$$

which can be further transformed to

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \mathbf{v} = 0 \,, \tag{2.16}$$

DOI:10.15774/PPKE.ITK.2020.007

taking into account Eq. (2.7). This is the linearized form of the well-known continuity equation.

The equation of state

The equation of state describes the relationship between the pressure and the density. It is clear that the pressure is a function of the density alone in inviscid[2] fluids. Besides, taking into account Eqs. (2.6) and (2.7) it can be expressed as

$$p_t\left(\rho_t\right). \tag{2.17}$$

To get the linearized form of the equation of state the Taylor expansion of the pressure $p_t\left(\rho_t\right)$ near the unperturbed state ρ_0 is used, and the higher order terms will be neglected because of the assumptions of Eqs. (2.6) and (2.7). Thus, we get the following form:

$$p_t = p_t\left(\rho_0\right) + \left(\rho_t - \rho_0\right) \frac{\partial p_t\left(\rho_t\right)}{\partial \rho_t}\Big|_{\rho_t=\rho_0}, \tag{2.18}$$

where $p_t\left(\rho_0\right) = p_0$ is the mean pressure at equilibrium, which results in

$$p_t - p_0 = \left(\rho_t - \rho_0\right) \frac{\partial p_t\left(\rho_t\right)}{\partial \rho_t}\Big|_{\rho_t=\rho_0}. \tag{2.19}$$

Substituting from Eqs. (2.6) and (2.7) causes

$$p = \rho \frac{\partial p_t\left(\rho_t\right)}{\partial \rho_t}\Big|_{\rho_t=\rho_0}. \tag{2.20}$$

The right-hand side multiplied by $\frac{\rho_0}{\rho_0}$ will be

$$\frac{\rho}{\rho_0}\, \rho_0\, \frac{\partial p_t\left(\rho_t\right)}{\partial \rho_t}\Big|_{\rho_t=\rho_0} = \frac{\rho}{\rho_0}\beta = \frac{\rho}{\kappa\rho_0}, \tag{2.21}$$

where

$$\beta = \rho_0\, \frac{\partial p_t\left(\rho_t\right)}{\partial \rho_t}\Big|_{\rho_t=\rho_0} \tag{2.22}$$

is the adiabatic bulk modulus by definition [72, p. 7], inverse of the compressibility κ. The final form of the equation of state can be written as follows:

$$p = \frac{\rho}{\kappa\rho_0}. \tag{2.23}$$

[2]inviscid fluid: ideal fluid without any viscosity

DOI:10.15774/PPKE.ITK.2020.007

Therefore, the equation of state means that the difference between the local and the mean pressure (differential pressure) is proportional to the deviation from the mean density (differential density) by some constant denoted as $\frac{1}{\kappa\rho_0}$ [71, p. 38].

2.2.3 The homogeneous linear wave equation

By combining the equation of state, the continuity equation and the equation of motion (Eqs. (2.24) to (2.26), respectively) it is possible to derive the homogeneous linear wave equation

$$p = \frac{\rho}{\kappa\rho_0}\,, \tag{2.24}$$

$$\frac{\partial\rho}{\partial t} + \rho_0\nabla\mathbf{v} = 0\,, \tag{2.25}$$

$$\rho_0\frac{\partial\mathbf{v}}{\partial t} = -\nabla p\,. \tag{2.26}$$

It is known that

$$\nabla^2 p = \nabla\cdot\nabla p\,. \tag{2.27}$$

Using it on Eq. (2.26) results in

$$\nabla\cdot\nabla p = \nabla\cdot\left(-\rho_0\,\frac{\partial\mathbf{v}}{\partial t}\right) = \frac{\partial}{\partial t}\left(-\rho_0\,\nabla\mathbf{v}\right). \tag{2.28}$$

The right-hand side of the equation appears in Eq. (2.25). After subtracting $-\rho_0\nabla\mathbf{v}$ from both sides and applying the differential operator $\frac{\partial}{\partial t}$ on both sides leads to:

$$\frac{\partial}{\partial t}\left(-\rho_0\,\nabla\mathbf{v}\right) = \frac{\partial^2\rho}{\partial t^2}\,. \tag{2.29}$$

Expressing ρ from Eq. (2.24) and using the differential operator $\frac{\partial}{\partial t^2}$ on both sides results in

$$\frac{\partial^2\rho}{\partial t^2} = \kappa\rho_0\frac{\partial^2 p}{\partial t^2}\,. \tag{2.30}$$

Substituting $\nabla\cdot\nabla p = \frac{\partial^2\rho}{\partial t^2}$, as they are equal according to Eqs. (2.28) and (2.29); furthermore, re-arranging everything to the left-hand side leads to an equivalent form of the homogeneous linear wave equation

$$\nabla^2 p - \kappa\rho_0\frac{\partial^2 p}{\partial t^2} = 0\,, \tag{2.31}$$

DOI:10.15774/PPKE.ITK.2020.007

which differs only in a constant from the well-known homogeneous linear wave equation used in acoustics. Thus, $\kappa\rho_0$ is equal to the inverse of speed of sound squared, from which

$$c^2 = \frac{1}{\kappa\rho_0}\,. \tag{2.32}$$

The left-hand side of Eq. (2.31) equals zero because no transformation between acoustic energy and heat is considered.

2.2.4 The inhomogeneous linear wave equation

So far, the media in which the US wave propagates, was considered to be homogeneous. However, most media is not uniform, in general it has a degree of inhomogeneity. Inhomogeneities cause scattering, and reflection will be at the point where is an interface between two media with different acoustic properties. In this section the wave equation for inhomogeneous media will be derived based mainly on [37, pp. 283–285].

Similarly to Eq. (2.7) it is assumed that the total density ρ_t and compressibility κ_t can be expressed as

$$\rho_t\left(\mathbf{r},t\right) = \rho_0 + \Delta\rho\left(\mathbf{r}\right) + \rho\left(\mathbf{r},t\right) = \rho_v\left(\mathbf{r}\right) + \rho\left(\mathbf{r},t\right) \tag{2.33}$$

$$\kappa_t\left(\mathbf{r}\right) = \kappa_0 + \Delta\kappa\left(\mathbf{r}\right) = \kappa_v\left(\mathbf{r}\right) \tag{2.34}$$

at any given location $\mathbf{r}$ and time t, where $\Delta\rho$ is the deviation from the mean density at equilibrium, $\Delta\kappa$ is the deviation from the mean compressibility at equilibrium, ρ_v and κ_v are the equilibrium density and compressibility, together with ρ, which is the small-signal acoustic density component caused by the propagating wave.

Using these expressions Eqs. (2.24) to (2.26) can be written as:

$$p = \frac{\rho}{\kappa_v\rho_v}\,, \tag{2.35}$$

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\left(\rho_v\mathbf{v}\right) = 0\,, \tag{2.36}$$

$$\rho_v\frac{\partial\mathbf{v}}{\partial t} = -\nabla p\,. \tag{2.37}$$

DOI:10.15774/PPKE.ITK.2020.007

By combining Eqs. (2.35) to (2.37) the same way as in Section 2.2.3, except that as the first step

$$\nabla \cdot \left(\frac{1}{\rho_v} \nabla p \right) \tag{2.38}$$

from Eq. (2.37) will be taken. The procedure results in the following form, similarly to Eq. (2.31):

$$\nabla \left(\frac{1}{\rho_v} \nabla p \right) - \kappa_v \frac{\partial^2 p}{\partial t^2} = 0 \,. \tag{2.39}$$

Let us take a look at the constant on the left-hand side, which gives

$$\frac{1}{\rho_v} = \frac{1}{\rho_0 + \Delta\rho} = \frac{\rho_0 + \Delta\rho}{(\rho_0 + \Delta\rho)^2} \,. \tag{2.40}$$

Assuming only small deviations from the mean density leads to

$$\frac{\rho_0 + \Delta\rho}{(\rho_0 + \Delta\rho)^2} \approx \frac{\rho_0 + \Delta\rho}{\rho_0^2} = \frac{1}{\rho_0} + \frac{\Delta\rho}{\rho_0^2} \,. \tag{2.41}$$

By substituting Eq. (2.41) to Eq. (2.39) results in

$$\nabla \cdot \left(\left[\frac{1}{\rho_0} + \frac{\Delta\rho}{\rho_0^2} \right] \nabla p \right) - \kappa_v \frac{\partial^2 p}{\partial t^2} = 0 \,, \tag{2.42}$$

which can be further rearranged into the following form:

$$\frac{1}{\rho_0} \nabla^2 p - \kappa_v \frac{\partial^2 p}{\partial t^2} = -\nabla \cdot \left(\frac{\Delta\rho}{\rho_0^2} \nabla p \right) . \tag{2.43}$$

At this point another substitution can be performed into the equilibrium compressibility κ_v using Eq. (2.34) and after some rearrangement it gives

$$\frac{1}{\rho_0} \nabla^2 p - \kappa_0 \frac{\partial^2 p}{\partial t^2} = -\nabla \cdot \left(\frac{\Delta\rho}{\rho_0^2} \nabla p \right) + \Delta\kappa \frac{\partial^2 p}{\partial t^2} \,. \tag{2.44}$$

To be able to use this formula in our context more transformation is needed. Let us multiply both sides by the mean density ρ_0

$$\nabla^2 p - \rho_0 \kappa_0 \frac{\partial^2 p}{\partial t^2} = -\nabla \cdot \left(\frac{\Delta\rho}{\rho_0} \nabla p \right) + \rho_0 \, \Delta\kappa \frac{\partial^2 p}{\partial t^2} \,, \tag{2.45}$$

and take the result mentioned at the end of Section 2.2.3 in Eq. (2.32) into account, which leads to

$$\nabla^2 p - \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = -\nabla \cdot \left(\frac{\Delta\rho}{\rho_0} \nabla p \right) + \frac{\Delta\kappa}{\kappa_0} \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2} = -\phi\left(\mathbf{r}, t\right) \,. \tag{2.46}$$

It can be noted that the left-hand side is the homogeneous linear wave equation, while the right-hand side is the expression of scattering caused by the local inhomogeneities of the medium (scattered field), denoted as $\phi\left(\mathbf{r}, t\right)$. This equation is called the inhomogeneous linear wave equation.

DOI:10.15774/PPKE.ITK.2020.007

2.2.5 Scattering

So far, there was no restriction on the pressure p. Taking scattering into account, the total pressure can be expressed as the sum of the incident and scattered pressure at any given location [34], denoted by p_i and p_s, respectively:

$$p = p_i + p_s \,. \tag{2.47}$$

The incident wave satisfies the homogeneous linear wave equation [34]:

$$\nabla^2 p_i - \frac{1}{c_0^2}\frac{\partial^2 p_i}{\partial t^2} = 0 \,. \tag{2.48}$$

Using this result an approximation can be made, namely:

$$\nabla^2 p_s - \frac{1}{c_0^2}\frac{\partial^2 p_s}{\partial t^2} = -\nabla \cdot \left(\frac{\Delta\rho}{\rho_0}\nabla p_i\right) + \frac{\Delta\kappa}{\kappa_0}\frac{1}{c_0^2}\frac{\partial^2 p_i}{\partial t^2} = -\phi\left(\mathbf{r}, t\right) \,, \tag{2.49}$$

which expresses that the scattered field s is generated at any given location $\mathbf{r}$ and any given time t only by the incident field . It is called *Born approximation* [37, pp. 287–289] [34, 38].

It can be shown that the pressure field in response to an acoustic disturbance at any given time and location (in the case of a non-zero source function and source volume) can be calculated using the solution of Green's function [71, p. 41] [72, p. 12]:

$$p\left(\mathbf{r}, t\right) = \int\limits_V \frac{\phi\left(\mathbf{r}', t - \frac{|\mathbf{r}-\mathbf{r}'|}{c}\right)}{4\pi\left|\mathbf{r}-\mathbf{r}'\right|} d\mathbf{r}' \,, \tag{2.50}$$

where $\mathbf{r}'$ denotes the location of the acoustical disturbance. The equation means that the pressure can be expressed by the spatial integral of all sources in the source region with appropriate time delays, and taking into account the spherical energy distribution during wave propagation.

As for the scattering strength, back-scattered energy in general is a result of the propagating wave either to be scattered or reflected. The former phenomenon describes an interaction with particles smaller than the wavelength, while in the case of the latter the particles are greater than the wavelength. Both physical phenomena occurs because of density inhomogeneity in the structure.

When the propagating mechanical wave reaches a boundary of two media having different densities, part of the energy travels through this boundary, while the

DOI:10.15774/PPKE.ITK.2020.007

remaining part will be reflected. In the case of 1-D scattering (180° reflection), the pressure impulse response can be expressed using one parameter only [37, p. 306], the acoustic impedance Z, and the ratio can be defined by the acoustic impedance difference of the two media. The acoustic impedance can be directly calculated using the density ρ and sound speed c of the medium as follows [37, p. 42]:

$$Z = \rho c \,, \tag{2.51}$$

and the reflection coefficient R is [37, p. 56]:

$$R = \left(\frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^2 , \tag{2.52}$$

where Z_1 is the acoustic impedance of the medium where the wave is travelling from, and Z_2 is where the wave is propagating to. The amplitude reflected back from the boundary is given by R, whereas the remaining $1 - R$ part of the incident amplitude is travelling through the boundary. The equation implicates that if the acoustic impedance difference $Z_2 - Z_1$ between the two media is great, the majority of the incident wave amplitude is going to be reflected; therefore, air (*e.g.*, lungs) or heavy tissue (*e.g.*, bone) is hard to image using US waves.

2.2.6 Complete shift-variant convolution model of US imaging

To describe the whole US image formation process (see Fig. 2.7), a setup given in Fig. 2.10 shall be considered. It is assumed that wave propagation is linear, non-attenuative, and non-dispersive, whereas variations in the acoustic properties of the medium should be small enough that only the incident wave is scattered (known as the Born approximation, see Section 2.2.5). Considering a single scatterer at $\mathbf{r}_0$, in order to obtain the output voltage the following equations can be formulated [37, p. 301] assuming a linear pulse-echo system (see Fig. 2.11):

$$p_r\left(\mathbf{r}_r, t\right) = e_i\left(t\right) * w_t^{\delta}(t) * \rho_0 \frac{\partial h_t\left(\mathbf{r}_0, t\right)}{\partial t} * h_s\left(\mathbf{r}_r, t\right) \,, \tag{2.53}$$

$$f(t) = 2 \int\limits_{S_r} p_r\left(\mathbf{r}_r, t\right) d^2\mathbf{r}_r \,, \tag{2.54}$$

DOI:10.15774/PPKE.ITK.2020.007

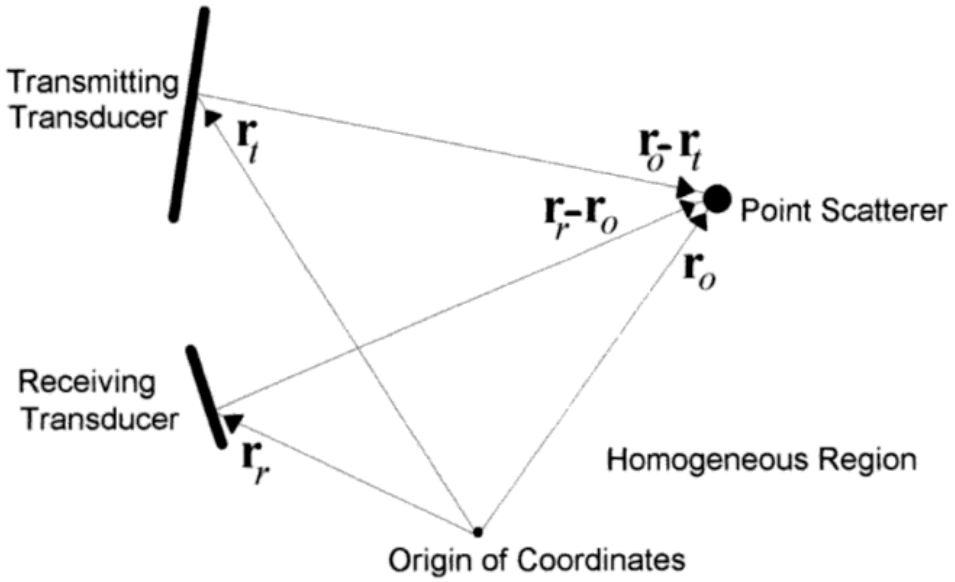

Figure 2.10: Geometry of a setup to determine the transmit-receive response of a system in the case of a single point scatterer. Picture taken from [37, p. 302].

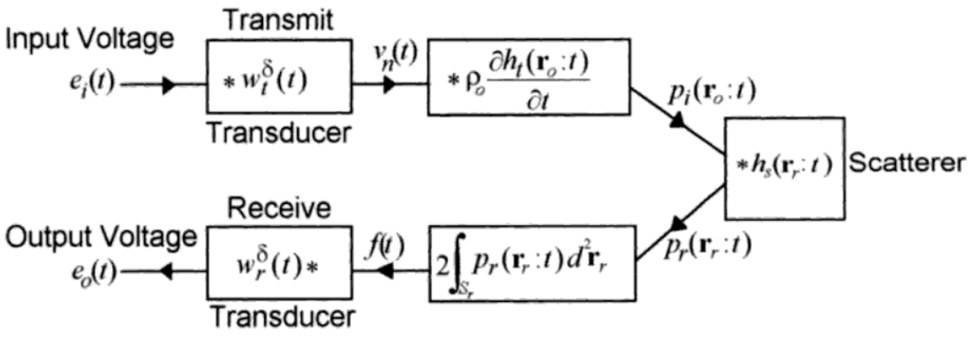

Figure 2.11: The transmit and receive responses in the case of a single point scatterer, assuming a non-attenuating medium. Picture taken from [37, p. 302].

$$e_0(t) = f(t) * w_r^{\delta}(t)\,, \tag{2.55}$$

where $p_r(\mathbf{r}_r, t)$ is the pressure distribution on the surface of the receiver transducer, $e_i(t)$ is the input voltage, $w_t^{\delta}(t)$ and $w_r^{\delta}(t)$ are the electromechanical responses of the transmit and receive transducers, respectively, $h_t(\mathbf{r}_o, t)$ is the velocity potential impulse response, $h_s(\mathbf{r}_r, t)$ is the pressure impulse response of the scatterer (the scattered pressure at an observation point caused by the incident wave), $f(t)$ is the total force acting on the surface of the receiver transducer, S_r and S_t are the area of the receive and transmit transducers, respectively. Substituting Eqs. (2.53)

DOI:10.15774/PPKE.ITK.2020.007

and (2.54) into Eq. (2.55) the output voltage is described as:

$$e_0(t) = \rho_0 e_i(t) * w_t^{\delta}(t) * \frac{\partial h_t(\mathbf{r}_0, t)}{\partial t} * \left[2 \int_{S_r} h_s(\mathbf{r}_r, t)\, d^2\mathbf{r}_r \right] * w_r^{\delta}(t)\,. \tag{2.56}$$

The impulse response of a point scatterer can be determined as follows:

$$h_s(\mathbf{r}_r, t) = s(t) * \frac{\delta\left(t - |\mathbf{r}_r - \mathbf{r}_0| / c_0\right)}{4\pi |\mathbf{r}_r - \mathbf{r}_0|}\,, \tag{2.57}$$

where $s(t)$ stands for the scatterer strength. As a consequence, the integral term in Eq. (2.56) can be be written as:

$$2 \int_{S_r} h_s(\mathbf{r}_r, t)\, d^2\mathbf{r}_r = s(t) * \int_{S_r} \frac{\delta\left(t - |\mathbf{r}_r - \mathbf{r}_0| / c_0\right)}{2\pi |\mathbf{r}_r - \mathbf{r}_0|} d^2\mathbf{r}_r = s(t) * h_t(\mathbf{r}_0, t)\ ; \tag{2.58}$$

therefore, Eq. (2.56) can be expressed as:

$$e_0(t) = \underbrace{\rho_0 \frac{\partial e_i(t)}{\partial t} * w_t^{\delta}(t) * h_t(\mathbf{r}_0, t) * h_t(\mathbf{r}_0, t) * w_r^{\delta}(t)}_{\text{PSF}} * \underbrace{s(t)}_{\text{SF}}\,, \tag{2.59}$$

which detailed form of the convolution-based image formation model was provided by Stepanishen [75], where the output voltage (which can be converted to pressure) can be calculated as the convolution between the point-spread function (PSF) of the system (which is the response of the imaging system to a point scatterer) and the SF.

2.2.7 Shift-invariant convolution model

The shift-invariant convolution model is a simplification of the shift-variant convolution model, which itself depends on several assumptions [34, 35] (see Section 2.2.6). Using such assumptions, it was shown that the RF image I can be estimated as the convolution of a SF with a PSF. Further assuming a spatially invariant PSF, meaning that the position of the scatterer relative to the US transducer is irrelevant in the terms of impulse of the scattering, leads to the shift-invariant convolution model:

$$I = \mathrm{SF} * \mathrm{PSF}\,, \tag{2.60}$$

where $*$ is the spatially invariant convolution operator. According to the Fourier theorem, Eq. (2.60) can be rewritten in the Fourier domain as

$$\mathcal{F}\{I\} = \mathcal{F}\{SF\}\, \mathcal{F}\{PSF\}\ , \tag{2.61}$$

DOI:10.15774/PPKE.ITK.2020.007

where $\mathcal{F}\{.\}$ represents the 2-D Fourier transform operator.

In the current work the imaging model described above was used for simulating the US images in Chapter 3.

2.3 Theory of resolution enhancement

2.3.1 Mathematical formulation

As it has been discussed in Section 2.2.7, an US image can be described as the convolution of the SF with the PSF of the imaging system. In general, the degradation process can be written as a convolution with a PSF, along with some added noise (often considered as zero mean Gaussian white noise) as follows:

$$y[i,j] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[k,l] \cdot x[i-k, j-l] + n[i,j], \tag{2.62}$$

where y is the observed image, x is the SF (the underlying structure), h is the PSF of the imaging system (alternatively, transfer function), and n stands for the noise. Equation (2.62) in a compact form reads as:

$$y = h * x + n, \tag{2.63}$$

where $*$ means the convolution operator. Equation (2.63) in the Fourier domain becomes:

$$Y = H.X + N, \tag{2.64}$$

where the capital letters denote the Fourier transforms of Y, H, X and N, respectively, and . stands for the element-wise multiplication. Estimating X seems straightforward in the absence of noise if the PSF is known, as it becomes a simple multiplication with the inverse of the Fourier spectrum of the PSF H^{-1}; however, there are many difficulties with this approach. From Eq. (2.64) it is obvious that separation of the $H.X$ product leads to an ill-posed problem, as it is not guaranteed that H^{-1} either exists or is unique, and noise also leads to inexact solutions. Furthermore, even if H^{-1} exists but the matrix is not well-conditioned, then regularization needs to be performed beforehand in order to limit the effect of noise (if H

DOI:10.15774/PPKE.ITK.2020.007

is small, then the noise will be greatly amplified). In US imaging H is band-limited, which makes the estimation of X more difficult, as resolution enhancement relies on recovering frequency components out of the frequency range of the data.

If the PSF is unknown, it becomes a blind deconvolution problem — which problem has been of long interest in research [76] —, where the PSF has to be estimated based on the data or a mathematical/physical model. An example is when the problem is transformed into the cepstral domain and the product of the two variables is separable, which imposes the importance of phase estimation of h and, consequently, the phase retrieval of x [77]. As a further consequence, phase retrieval algorithms have been developed [78–80].

Primarily, there are two ways to estimate the original image x from the observed image y: one either have to estimate h separately, or h and x jointly. As for the first one, it becomes a classical deconvolution problem, while the second one in general means an alternating minimization of a cost function [81, 82].

The following three subsections (Sections 2.3.2 to 2.3.4) will be about the separate estimation of the PSF, the classical deconvolution methods and joint estimation of the PSF and the SF.

2.3.2 PSF estimation

As it has been mentioned previously in Section 2.3.1, one way to assess the SF is to estimate the PSF separately and perform deconvolution.

To estimate the PSF some assumptions has to be made. Namely, that it acts as a low-pass filter, having a slowly varying amplitude and phase. From this, the following can be done: calculating the Fourier spectrum of the whole image (or a large part) and setting its phase to zero. The problem with this approach is that it violates causality, which would be a reasonable assumption in the axial direction taking the forth- and back-propagation of the US waves into account.

It is possible to construct a causal signal from any 1-D signal by making it minimum phase. This ensures that the highest energy concentration of the signal is as close to $t = 0$ as possible and $h(t) = 0$ for all $t \leq 0$. A minimum-phase estimate $h_{mp}(t)$ may be obtained from the 1-D signal $h(t)$ (which can mean an A-line in US

DOI:10.15774/PPKE.ITK.2020.007

imaging) using processing in the cepstral domain [83][84, p. 998]:

$$H(f) \leftarrow \mathcal{F}\{h(t)\}\,, \tag{2.65}$$

$$C(t) \leftarrow \log(|H(f)|)\,, \tag{2.66}$$

$$C_{mp}(t) \leftarrow \begin{cases} 0 & t < 0 \\ C(t) & t = 0 \\ 2C(t) & t > 0 \end{cases}\,, \tag{2.67}$$

$$h_{mp}(f) \leftarrow \exp\left(\mathcal{F}\left\{C_{mp}(t)\right\}\right)\,, \tag{2.68}$$

$$h_{mp}(t) \leftarrow \mathcal{F}^{-1}\left\{H_{mp}(f)\right\}, \tag{2.69}$$

where $\mathcal{F}\{.\}$, $\mathcal{F}^{-1}\{.\}$ operators mean the Fourier and inverse Fourier transform, respectively. Note that typically a set of absolute Fourier spectra $|H(f)|$ are averaged and used during the log-transform step (Eq. (2.66)).

This minimum phase approach was used in the literature [83], where an axial deconvolution model was considered. It also implies that the (de)convolution process is separabale into axial, transversal and elevational directions. However, there seems to be no mutual agreement on the validity of the minimum phase assumption. There was experimental evidence for confirming it provided by Adam and Michailovich [85]. Nonetheless, Taxt reported in a review [86] that the minimum phase approach yielded the worst results.

There are additional techniques to estimate the PSF, such as an autoregressive moving average (ARMA) model used by Jensen [87]. Interestingly, Jirik *et al.* [88] reports it is possibly not suitable for US imaging, as the ARMA model inherently assumes fairly smooth original images. Additionally, using higher order statistics it is also a possibility to estimate the cepstrum of the axial PSF [89], moreover, it can also be extended to several dimensions as reported by Wan *et al.* [90].

Instead of the minimum phase approach it is also possible to use homomorphic filtering, which is about separating the PSF and SF components in cepstral space by using Butterworth filtering. As it was mentioned, the PSF is assumed to be slowly varying compared to the SF, which provides the basis for separation. Utilizing this, and the fact that in cepstral space Eq. (2.64) can be written as [91, 92]:

$$\log(Y) = \log(H) + \log(X) = \log|H| + \log|X| + j\measuredangle(H) + j\measuredangle(X)\,, \tag{2.70}$$

DOI:10.15774/PPKE.ITK.2020.007

whereas the noise term is ignored for simplicity and $\measuredangle$ stands for the phase of the complex functions.

2.3.3 Classical deconvolution-based methods

This section is about the most widely used deconvolution methods, namely the Wiener and Lucy–Richardson deconvolution techniques. As there are many different classical deconvolution approaches to solve image restoration or enhancement problems, for a more comprehensive review the reader is directed to [93, Chapter 3].

Wiener deconvolution

Wiener deconvolution (often referred to as Wiener filtering) offers an optimal solution in the least square sense [94–96]. From the degradation model Eq. (2.63), we would like to estimate the original unknown input x as accurately as possible. In mathematical terms, we wish to solve the following:

$$\hat{x} = \arg\min_{\hat{x}} \mathbb{E}[x - \hat{x}]^2 \,, \tag{2.71}$$

where the estimated input is denoted as $\hat{x}$. To do so, we look for the optimal filter g which can recover $\hat{x}$ from the observed signal y as follows:

$$\hat{x} = g * y \,, \tag{2.72}$$

where $*$ means the convolution operator. Eq. (2.72) in the Fourier domain becomes

$$\hat{X} = G \cdot Y \,, \tag{2.73}$$

where the capital letters $\hat{X}$, G and Y stand for the Fourier transforms of $\hat{x}$, r and y, respectively.

The mean squared error (MSE) is an expectation value in Eq. (2.71); therefore, it is enough to know (or more precisely assume) the statistical characterization of the original input signal defined as $\hat{x}$. It can be shown that the optimal filter which should be used to minimize Eq. (2.71) is as follows [96, Chapter 11]:

$$\hat{X} = \frac{H^*}{|H|^2 + S_N/S_X} \cdot Y \,, \tag{2.74}$$

DOI:10.15774/PPKE.ITK.2020.007

where H^* is the complex conjugate of the Fourier spectrum of the PSF h, S_N and S_X mean the power spectrum density (magnitude of the Fourier transform of the correlation function) of the noise and signal, respectively. Note that in many approaches the ratio S_N/S_X is often considered as a constant, which simplifies the whole estimation of $\hat{X}$. If Gaussian zero mean white noise is assumed (as it has already been mentioned in Section 2.3.1), Eq. (2.74) is simplified to:

$$\hat{X} = \frac{H^*}{|H|^2 + \sigma^2/S_X} \cdot Y \,, \tag{2.75}$$

where σ is the variance of the noise, and the ratio σ^2/S_X means the noise-to-signal ratio (NSR). By assuming no additive noise, the Wiener filter becomes the ideal inverse filter.

Lucy–Richardson deconvolution

As it has been shown in the previous section, the Wiener deconvolution is an optimal minimum MSE estimator algorithm in the least square sense, and is more practical to use in the case of Gaussian noise. While both the Wiener and Lucy–Richardson deconvolution methods are minimum mean square estimators, in contrast to Wiener deconvolution, in the case of Poisson noise the iterative Lucy–Richardson method [97–99] can be used. Poisson noise is usually assumed when using photo diodes or the so-called CCD sensors for capturing images (however, promising results can also be found in the case of US [100]); therefore, this technique is mostly used for astronomical images [57,101–103], where it is often considered as the golden standard to which new techniques are compared. It is often called as the expectation-maximization algorithm, and has no closed form solution as it is an iterative approach. It can be described as follows [99]:

$$\hat{x}_{k+1} = \hat{x}_k \left(h \otimes \frac{y}{h * \hat{x}_k} \right) \triangleq \psi\left(\hat{x}_k\right) \,, \tag{2.76}$$

where $*$ and $\otimes$ are the convolution and correlation operators, respectively, $\hat{x}_k$ stands for the estimate of x after k number of iterations, and $\psi\left(.\right)$ is the Lucy–Richardson function. The algorithm usually suffers from noise amplification if the convergence criteria is not well defined, which is one of the drawbacks of maximum likelihood

DOI:10.15774/PPKE.ITK.2020.007

techniques. In general, a single pixel changes rapidly only during the first few iterations of a restoration process, and the convergence can occur slowly or even in an oscillatory way. In the lights of this, the running time is of great importance for large data sets; thus, there are many ways to reduce the computational time [104–109].

2.3.4 Cost function minimization

As it has been mentioned previously described in Section 2.3.1, one way of dealing with the problem of blind deconvolution is to jointly estimate the PSF and the SF. To this end, one can perform a Bayesian formulation, where the deconvolution algorithm is based on statistical assumptions about the SF, the PSF, and the noise. It is solved as an alternating minimization problem [88], which is an extension of the homomorphic deconvolution. It is also possible to use stochastic sampling with a few pre-selected hyperparameters, which can provide a larger solution space [110, 111]. Minimizing a cost function J, which contains both h and the residual, essentially defines what type of solutions are preferred:

$$J = \|y - h * x\|_2^2 + \Lambda \|x\|_\gamma \,, \tag{2.77}$$

where Λ is an appropriate weight, and γ is the p-value for the l_p norm, and $0 \leq p \leq 2$. Values closer to 0 mean a more sparse restriction. Note that not only l_p norm can be used (*e.g.*, total variation (TV) [112, 113]). The data fidelity term $\|y - h * x\|_2$ ensures that the measured image y, and the convolution between the reconstructed image and the PSF stay as close as possible, while the $\|.\|^2$ assures a smooth function (which simplifies the minimization), and $\Lambda\|x\|_\gamma$ provides the desired sparsity. To reduce the optimization time, certain parameterized PSFs can be used (*e.g.*, Gaussian-modulated sinusoids [110]). It is also possible to perform axial deconvolution only [114, 115] while projecting the signal into a low-resolution (envelope) space, and performing alternating direction method of multipliers (ADMM). However, even if the PSF and the SF are estimated jointly, a first estimate of the PSF (see Section 2.3.2) is essential [88, 112].

DOI:10.15774/PPKE.ITK.2020.007

2.3.5 Deep learning

Artificial intelligence and its sub-fields, machine learning (ML) and deep learning (DL), are a subject of interest in ongoing research, where computers can perform different tasks such as automatically recognizing complex patterns in radiology [116] or cardiology [117], automating text categorization [118], learning to play complex games like Go [119], or even countering social media bots manipulating public opinion [120] by learning from previous experiences and therefore evolve, similarly to humans. Specifically, DL witnessed a booming interest since 2012, when the ImageNet challenge has been won by a deep neural network architecture, the AlexNet [121]. Such an architecture consist of artificial neurons (Fig. 2.12) forming multi-layered

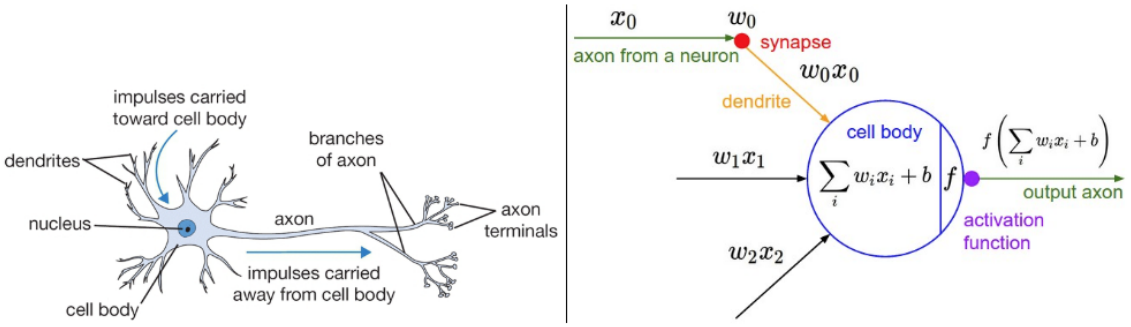

Figure 2.12: Schematic of a neuron (left) and its corresponding mathematical model (right). Picture taken from [122].

artificial neural networks (ANN). Each neuron performs a dot product of the input x_0 from the dendrites and its weight(s) w_0, plus adds the bias b. Then, the activation function $f(z)$ is being used on this result and is forwarded to the next layer as an output (see Fig. 2.13). Each layer performs a different operation to extract certain features of the input, and uses the output of the previous layer. Upon completion, it provides its own output to the next layer, similarly to the logical process of human neural system. The main difference between ML and DL is while the latter one can essentially learn and improve based on the difference between the actual output and the desired output, the previous one is not able to evolve without human intervention.

The output of the first layer (see Fig. 2.13) can be calculated with simple matrix

DOI:10.15774/PPKE.ITK.2020.007

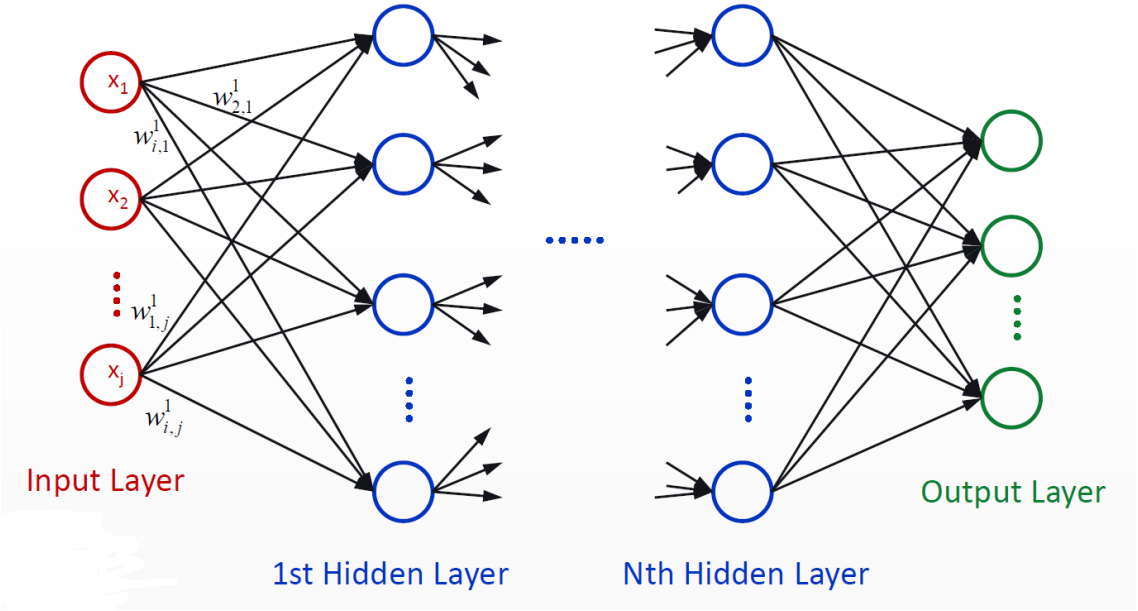

Figure 2.13: Network layers. Picture taken from [123].

operations as follows:

$$\begin{bmatrix} a_1^1 \\ a_2^1 \\ \vdots \\ a_m^1 \end{bmatrix} = f\left(\begin{bmatrix} z_1^1 \\ z_2^1 \\ \vdots \\ z_m^1 \end{bmatrix}\right) = f\left(\begin{bmatrix} - & w_1^1 & - \\ - & w_2^1 & - \\ & \vdots & \\ - & w_m^1 & - \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1^1 \\ b_2^1 \\ \vdots \\ b_m^1 \end{bmatrix}\right), \tag{2.78}$$

which in a compact form reads as:

$$a^{(1)} = f\left(z^{(1)}\right) = f\left(W^{(1)}x + b^{(1)}\right). \tag{2.79}$$

The in- and outputs of the further layers (shown until the 3rd) can be calculated using the chain rule:

$$\begin{aligned} a^{(1)} &= f\left(z^{(1)}\right) = f\left(W^{(1)}x + b^{(1)}\right), \\ a^{(2)} &= f\left(W^{(2)}a^{(1)} + b^{(2)}\right), \\ z^{(3)} &= W^{(3)}a^{(2)} + b^{(3)}. \end{aligned} \tag{2.80}$$

In the current work convolutional neural networks (CNN) are used (see Chapter 5). Regular fully connected layers, where all units of a layer are connected to all units of the next layer, do not scale well to images. Considering an ANN with

DOI:10.15774/PPKE.ITK.2020.007

input images having a size of 1 MegaPixel (1024×1024 pixels), if only 100 neurons are in the first layer that yields more than 100 million weight parameters already. To precisely train that many parameters one would need a huge data set to avoid overfitting, not mentioning the greatly increased computational cost.

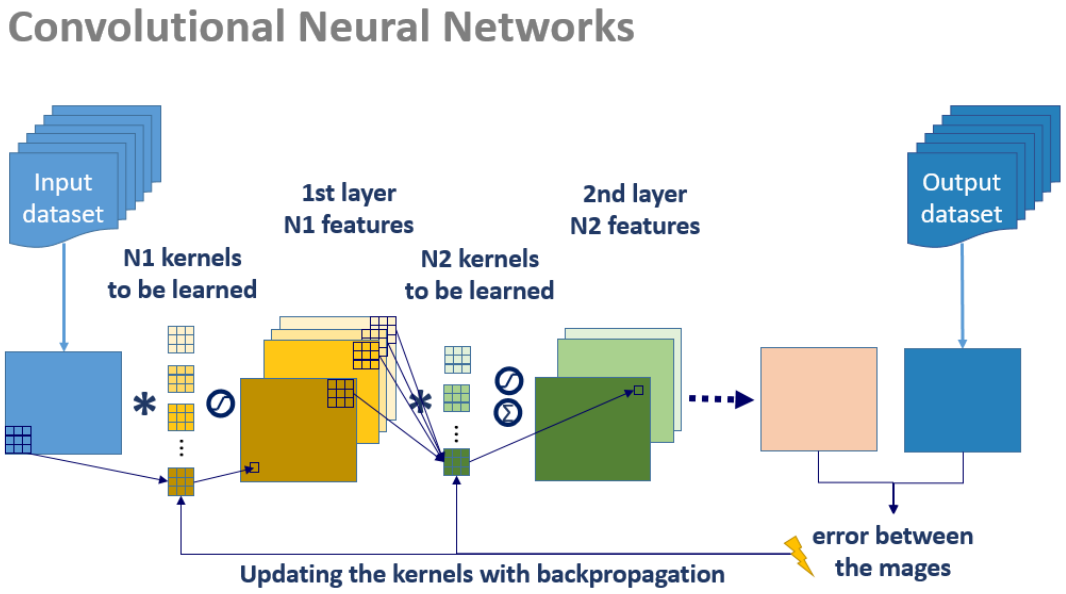

Figure 2.14: Convolutional Neural Network. Image taken from [124].

CNNs take advantage of the fact that local correlation and recurring local statistics are typical for general images; therefore, the same weights are used for different parts of the image. Figure 2.14 shows a typical CNN. First, N1 different number of convolutional kernels are going to be applied on the input image. After the kernel is applied, a (derivable) non-linear activation function is used on the result. This operation yields one pixel in the first image in the first layer. Next, the kernel is shifted in the input image in order to calculate the next pixel, which continues until the last one is calculated. N1 kernels yield N1 images in the first layer. Next, all the images in the first layer act as input images, on which the N2 different kernels will be applied. The average of N1 images convolved with a given kernel will be one of the N2 resulting images in the second layer. The same procedure is repeated until the last layer. Upon completion, the output is compared to the reference output, and the error is measured. Using this quantified error measure it is possible to update the kernels with backpropagation, thus the CNN can learn from the previous

DOI:10.15774/PPKE.ITK.2020.007

experiences and evolve. This way, for the first layer only $N1 \cdot S_{kernel}$ weights need to be learned, where S_{kernel} is the size of the kernel.

It is also possible to use pooling (effectively downsampling), which keeps usually only the largest or average value of the local neighbourhood. It helps to reduce the dimensionality of the next layer, resulting in less parameters to train.

2.3.6 US specific methods

Equivalent scatterers

As it was discussed in Section 2.2.5, every time an ultrasound examination takes place a mechanical wave insonifies the examined media, where the waves get reflected/scattered due to the structure's inhomogeneity (in density [34], acoustic impedance [35] or bulk modulus [38]). The backscattered pulses will interfere, which can be either constructive or destructive [125] with respect to their relative phase, resulting in a signal with speckle pattern. However, this statement immediately raises some questions: is the so-called backscattered signal unique? Are there other scatterer distributions (possibly containing fewer scatterers), which could lead to the same result? Furthermore, if such a function – depicting equivalent scatterers – exists, how much of a reduction in computational time could this achieve in simulations? Dantas *et al.* [125] introduced an elegant method to tackle this question. However, it is mainly based on processing steps made in the frequency domain. The purpose of the current section is to consider the equivalent steps in the spatial domain in order to better understand the problem of equivalent scatterers.

As it was previously shown in Section 2.2.7, every linear system can be described by performing convolution between the desired impulse response and the transfer function of the system. In the case of US signals this implies that a so-called A-line $i(z)$ — depicting the backscattered signal — restricted to the axial direction z can be described as (see Eq. (2.60)):

$$i(z) = s(z) * p(z) \ , \tag{2.81}$$

where $s(z)$ is the SF, $p(z)$ stands for the PSF and $*$ denotes the convolution operator. In the frequency domain, this convolution operator becomes a multiplication

DOI:10.15774/PPKE.ITK.2020.007

of the corresponding spectra (see Eq. (2.61)):

$$I(u) = S(u) \cdot P(u) \ , \tag{2.82}$$

where u is the angular frequency in the axial direction, and $I(u)$, $S(u)$ and $P(u)$ are the Fourier transformations of $i(z)$, $s(z)$ and $p(z)$, respectively.

The aforementioned Fourier-based equivalent scatterer method [125] will now be described. An impulse response with a central frequency u_0 and bandwidth BW causes the backscattered signal to be band-limited, which effectively yields a windowing on the original SF as:

$$S_{lim}(u) = S(u) \cdot \Pi\left(\frac{u - u_0}{BW}\right) \ , \tag{2.83}$$

where $\Pi(.)$ denotes the rectangular function, and $S_{lim}(u)$ is the band-limited SF. Afterwards, this effective bandwidth is repeated in the frequency domain by performing a convolution with an appropriate comb function:

$$S_{samp}(u) = S_{lim}(u) * \mathrm{III}\left(\frac{u}{BW} - u_0\right) \ , \tag{2.84}$$

where $\mathrm{III}(.)$ is the Dirac comb (or the so-called Shah) function. In the spatial domain this will result in a new set of scatterers. When this new SF is getting convolved with the impulse response, the result is an A-line that is characteristically equivalent to the original one:

$$i_{samp}(z) = s_{samp}(z) * p(z) \ , \tag{2.85}$$

$$i_{samp}(z) \approx i(z) \ . \tag{2.86}$$

In the following these steps are realized in the spatial domain. Multiplication with a rectangular function in the frequency domain means convolution in the spatial domain of Eq. (2.83):

$$\begin{aligned} s_{lim}(z) &= \mathcal{F}^{-1}\left\{S(u) \cdot \Pi\left(\frac{u - u_0}{BW}\right)\right\} = \\ &= s(z) * \frac{e^{i\left(u_0 + \frac{BW}{2}\right)z} - e^{i\left(u_0 - \frac{BW}{2}\right)z}}{iz} \ , \end{aligned} \tag{2.87}$$

where

$$\frac{e^{i\left(u_0 + \frac{BW}{2}\right)z} - e^{i\left(u_0 - \frac{BW}{2}\right)z}}{iz} = 2 \cdot e^{iu_o z} \cdot \mathrm{sinc}\left(\frac{BW}{2} z\right) \ . \tag{2.88}$$

DOI:10.15774/PPKE.ITK.2020.007

It can be seen that the SF is convolved by a sinc function, which usually represents the first step in a downsampling operation. The frequency of the sinc function will directly determine the quality of the downsampling: the wider the bandwidth, the smoother the resulting resampled function will be. Next, Eq. (2.84) can be written in the spatial domain as follows:

$$\begin{aligned} s_{samp}(z) &= \mathcal{F}^{-1}\left\{ S_{lim}(u) * \mathrm{III}\left(\frac{u}{BW} - u_0\right)\right\} = \\ &= s_{lim}(z) \cdot \mathrm{III}(z \cdot BW) \cdot e^{-iu_o z}, \end{aligned} \tag{2.89}$$

where the multiplication with a Dirac comb represents sampling. Performing a convolution with the impulse response function:

$$\hat{i}_{samp}(z) = s_{samp}(z) * p(z), \tag{2.90}$$

which should result in

$$\hat{i}_{samp}(z) \approx i(z). \tag{2.91}$$

Note that the approach above is similar to the one suggested by Mo *et al.* [126], where the effective scatterer of each voxel position is calculated as the average of all scatterers in the voxel. Here the difference is that an exact mathematical formulation is provided for how the averaging should take place for accurate results.

Recently, our group has also shown that with a stippling based algorithm arbitrary medical US images can be well approximated (having $R^2 \geq 80\%$) and physically realized. For more details the reader is directed to [Au6] [127].

Axial processing

One set of techniques of interest fall under the category of axial filtering, which have recently been shown by our group to improve lateral resolution [Th2]. Since they are relatively easy to implement and do not require an estimate of the full (and shift-variant) 3-D response, they are potentially well-suited to single-element imaging systems such as SAM, whose response is highly depth-dependent.

From Eq. (2.60), the spatially variant PSF can be further decomposed into two components [34, 38] [Th1, Th2], whereas only one of them is spatially variant [83, 128] [Th2]:

$$PSF = v_{pe} * h_{pe}, \tag{2.92}$$

DOI:10.15774/PPKE.ITK.2020.007

where v_{pe} is the so-called electric response of the ultrasound transducer transformed into the spatial domain (note that this also includes the electrical excitation of the transducer as well as the pulse-echo response); and h_{pe}, the so-called geometric response, which is the pulse-echo spatial impulse response of the transducer, including the geometric focusing and beam-forming properties. Whereas the 2-D geometric response varies spatially, the electric response is spatially invariant and acts exclusively along the axial dimension. This means that shift-invariant axial filtering of the image has the potential to deconvolve part of the PSF out of the image. Paradoxically, although the filtering acts in the axial direction, this can also improve lateral resolution [Th2], as discussed in the following. For simplicity, in the following we refer to estimating the 1-D PSF of the imager using various techniques, assuming that this estimate will be dominated by the shift-invariant electric response v_{pe}, as well as possibly including a shift-invariant component of the geometric response h_{pe}.

To consider how axial filtering of the signal – including 1-D deconvolution – affects the US image, we note that the lateral and axial resolutions R_L, R_A of a single-element transducer can be estimated using the following formulas [37, p. 173 and 511, respectively]:

$$R_L = 1.22 \cdot f_{\#} \frac{c}{f_0}\,, \tag{2.93}$$

$$R_A = \frac{c}{2 \cdot BW}\,, \tag{2.94}$$

where c stands for the speed of the propagating wave in the medium, BW for the bandwidth, $f_{\#}$ for the f-number and f_0 for the central frequency of the transducer. Note that Eq. (2.94) is an equivalent form of Eq. (2.4), and Eq. (2.93) is the same as Eq. (2.3). Considering Eq. (2.94), the axial resolution depends solely on the bandwidth (from the point of view of the imaging system). Hence, an increase in bandwidth yields a better axial resolution. Therefore, performing depth-independent axial deconvolution can provide an increase in axial resolution. Secondly, if the bandwidth is increased more at the upper cut-off frequency, or higher frequencies are weighted more, the central frequency of the signal is also increased, providing the possibility to simultaneously enhance the resolution in both axial and lateral directions, given the fact that the high frequency components of the PSF are con-

DOI:10.15774/PPKE.ITK.2020.007

centrated in its central part (see Fig. 2.15), and their amplification would result in a laterally increased resolution. However, the amplification of high frequencies comes at the risk of a loss in signal-to-noise ratio (SNR).

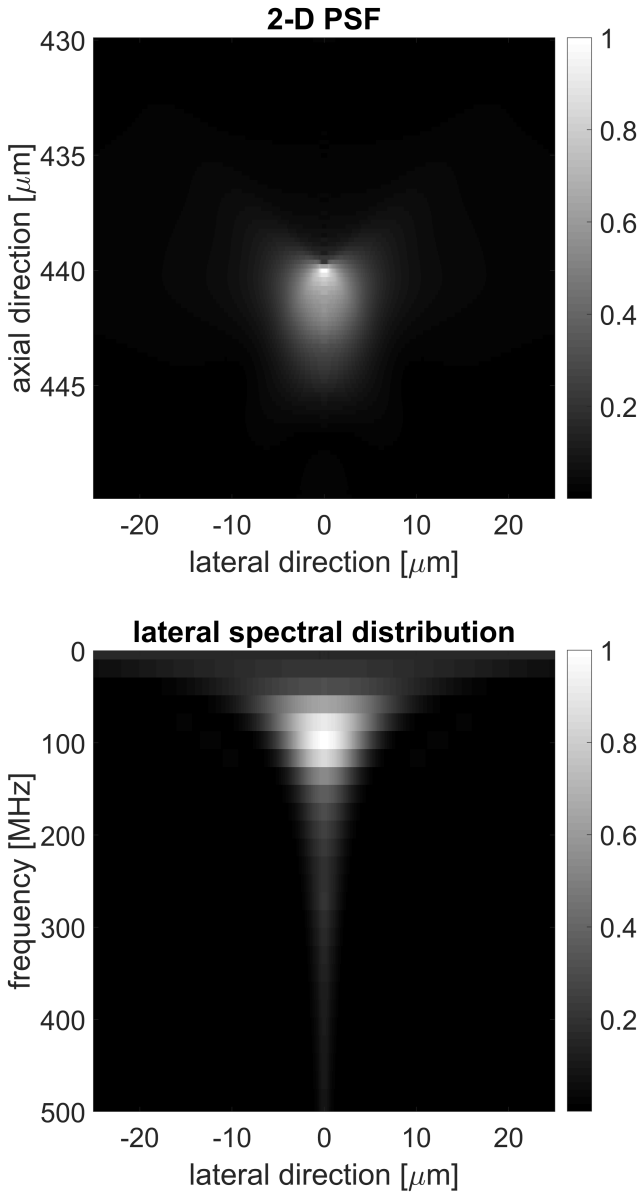

Figure 2.15: Field II-simulated 2-D PSF (top) of a single-element transducer (187 MHz center frequency, 440 µm focal distance) and its corresponding lateral spectral distribution (bottom). Laterally the farther is a point from the central part of the PSF the lower frequency components can be found.

Based on the above, it is hypothesized that suitable axial filtering of the data can be used to improve lateral resolution in SAM images, as already demonstrated

DOI:10.15774/PPKE.ITK.2020.007

for US images [Th2]. One possible approach is to perform 1-D deconvolution using an estimate of the 1-D PSF. One method of obtaining such an estimate is to average the Fourier spectra of the A-lines and set the phase to zero, resulting in a zero-phase estimate (ZP). Alternatively, and as is more common in the ultrasound literature given the causality of the electric response, a minimum-phase estimate h_{mp} may be obtained using processing in the cepstral domain, as it has been previously shown in Section 2.3.2.

If the mentioned estimate is obtained, an axial deconvolution can be performed on every A-line independently.

DOI:10.15774/PPKE.ITK.2020.007

Chapter 3

Experimental validation of ultrasound image formation

3.1 Introduction

Models of ultrasound image formation describe the forward process of how an ultrasound image is formed from an acoustic medium. Such models can be used to generate simulated ultrasound images or to obtain quantitative descriptors of the medium from real ultrasound images. A relatively simple and widely used model of image formation treats the ultrasound image (before envelope detection and compression) as the shift-invariant convolution of the imaging system point spread function (PSF) with the scattering function (SF) of the medium [129] (see also Section 2.2.7). In the following paragraphs, the uses of this model are briefly reviewed to provide the context for the current work.

In one of the earlier applications of the shift-invariant convolution model, Bamber and Dickinson [40] generated random SFs by assuming an exponential or a Gaussian autocorrelation (AC) function. Each SF was convolved with a Gaussian-modulated sinusoid PSF, and the resulting ultrasound images were found to be qualitatively similar to real ultrasound images of a sponge. The simulations were also able to predict a so-called banding feature visible in ultrasound images in cases where the correlation length of the SF was shorter than one of the beam dimensions. Other authors employed different SF models to generate ultrasound images. Using a so-

DOI:10.15774/PPKE.ITK.2020.007

called marked regularity model to place point scatterers in the SF, Cramblitt and Parker [130] convolved the SFs with a gated cosine PSF and found that a wide range of envelope distributions (Rayleigh, Rician, homodyned K) could be generated that are commonly observed in experimental data. Jacobs and Thijssen [131] used structural as well as diffuse components in the SF to make predictions about first- and second-order statistics that would be observed from ultrasound images of the liver.

The shift-invariant convolution model has also been used to simulate ultrasound images using biologically derived SFs. Waag *et al.* [132] used images of pig liver lobules to investigate the relationship between backscattered spectra in 1-D, 2-D, and 3-D. Lizzi *et al.* [133] used acoustic microscope data to estimate speed of sound maps of tissue, which were used to generate simulations of conventional frequency ultrasound images. The use of histology to estimate SF is also receiving increased attention: Mamou *et al.* [39] generated 3-D acoustic impedance maps from 3-D volumes of H&E-stained histology, which was then used to estimate quantitative acoustic parameters [134]; Daoud and Lacefield [135] used histology images to derive spatial statistics of the SF. Gyöngy *et al.* [136] performed simulations using SFs derived from histology images and compared the first-order statistics of these images with those of corresponding real ultrasound images.

Increasingly, there has been a trend toward generating fast simulations of ultrasound images. Considering the issue of computational speed, shift-invariant convolution is an appropriate choice because the operation is equivalent to multiplication in the Fourier domain. In the work of Hergum *et al.* [41], cardiac images were simulated and their appearance was found to be qualitatively similar to images simulated using a shift-variant convolution model. Gao *et al.* [137] also simulated cardiac images but with shift-variance in the axial direction. Bürge *et al.* [138] combined shift-invariance with ray tracing to generate simulated images of phantoms and tissue volumes.

As can be observed from the above representative review, the shift-invariant convolution model is widely used in the literature yet the comparison of the resulting simulations with experiments is either qualitative or based on aggregate descriptors such as envelope statistics or spectral components. This is due to the lack of an

DOI:10.15774/PPKE.ITK.2020.007

experimental setup where the scattering distribution of an acoustic medium and the corresponding ultrasound image can be simultaneously recorded. In the current work, macrophotographs of micrometer-scale spherical scatterers were used to derive the SF of a medium that was also simultaneously imaged with an ultrasound system. Using the shift-invariant convolution model and estimates of the PSF, the one-to-one spatial correspondence between ultrasound simulations and experiment was quantitatively compared. Such quantitative comparisons allowed a validation of the shift-invariant convolution model, first by investigating the best simulation accuracy using different estimates of the SF and PSF, and then by considering the extent of different sources of error in the simulation.

3.2 Methods

Section 2.2.7 describes how an ultrasound image can be modeled as the shift-invariant convolution of a SF and PSF. Then, the theoretical background for estimating the SF and PSF is given, so that the model can be used to generate simulated ultrasound images. Next, the coefficient of determination R^2 is described, which allows the accuracy of the ultrasound simulations to be evaluated. Finally, possible sources of error between the real and simulated ultrasound images are provided. These theoretical considerations provide the basis for the research methods described in the next section.

3.2.1 SF estimation

The SF originates from deviations of the acoustic properties of the medium from some mean. Based on differing assumptions, various expressions exist for SF: for instance, as a combination of fractional changes in density and speed of sound [34]; as fractional changes in bulk modulus [38]; or as fractional changes in acoustic impedance [35]. The following estimates were used:

1. threshold-based: in the case of homogeneous scatterers placed in a homogeneous background, it is correct (up to a scaling factor) to set the estimated

DOI:10.15774/PPKE.ITK.2020.007

scattering function $\mathrm{SF}_{\mathrm{est}}$ to 1 where there is a scatterer, and set it to 0 where there is a background [36]

2. projection-based: for a spherical scatterer in the plane of a linear array, $\mathrm{SF}_{\mathrm{est}}$ may be treated as a circular cross-section or as the projection of the sphere function onto the 2-D plane (similarly to [36, 132])

3. point-based: alternatively, for acoustically compact spherical scatterers, $\mathrm{SF}_{\mathrm{est}}$ may be set to a collection of discrete points.

These estimates will later be referred to as threshold-based, projection-based, and point-based estimates, respectively.

3.2.2 PSF estimation

The PSF of the imaging system arises from the combination of the incident pressure field transmitted by the transducer and the spatiotemporal response of the transducer and receive beamformer to the scattered field. The following estimates were used:

1. Field II-based: one method of estimating PSF [137] is using the spatial impulse response approach at a particular depth [34], using a package such as Field II [128, 139]

2. data-based: alternatively, for a region where only one point scatterer is present, SF may be treated as a Dirac delta function. Thus, if the resulting image I_{real} is available, then according to Eq. (2.60), PSF may be estimated from it

3. deconvolution-based: Lastly, if SF consists of a distribution of scatterers, then using Eq. (2.61), an estimate $\mathrm{PSF}_{\mathrm{est}}$ can be obtained as follows:

$$\mathrm{PSF}_{\mathrm{est}} = \mathcal{F}^{-1}\left\{\frac{\mathcal{F}\left\{I_{\mathrm{real}}\right\}}{\mathcal{F}\left\{\mathrm{SF}_{\mathrm{est}}\right\}}\right\}, \tag{3.1}$$

where $\mathcal{F}^{-1}\{.\}$ denotes the inverse Fourier transform.

The above estimates will later be referred to as Field II-based, data-based, and deconvolution-based estimates, respectively.

DOI:10.15774/PPKE.ITK.2020.007

3.2.3 Evaluation of simulation accuracy

When generating a simulated image I_{sim} from estimates SF_{est} and PSF_{est}, I_{sim}, and the real ultrasound image I_{real} will differ by an additive error term E:

$$I_{\text{real}} = I_{\text{sim}} + E\,, \tag{3.2}$$

$$I_{\text{sim}} = a\text{SF}_{\text{est}} * \text{PSF}_{\text{est}}\,, \tag{3.3}$$

where a is an unknown scaling term arising from the scale-unknown estimates SF_{est} and PSF_{est}; this term needs to be estimated using linear regression. Due to the partitioning of variance [140, pp. 21–23], the coefficient of determination R_I^2 can be defined as [140, pp. 92, 166]

$$R_I^2 = 1 - \frac{\sigma_E^2}{\sigma_{I_{\text{real}}}^2} = \frac{a^2\sigma_{I_{\text{sim}}}^2}{\sigma_{I_{\text{real}}}^2}\,, \tag{3.4}$$

where σ_u^2 denotes the variance of all the pixel values of image u. The above equation leads to an interpretation of R_I^2 as the proportion of variance (or signal power) in the real ultrasound image that can be explained by the simulated image [141]. In a similar vein, R_B^2 considers the proportion of variance in the real B-mode image B_{real} that is accounted for by the simulated B-mode image B_{sim}. One method of computing the coefficient of determination is taking the square of the Pearson correlation coefficient [141].

3.2.4 Sources of error

There can be several sources of simulation error E. First, the estimates SF_{est} and PSF_{est} may be inaccurate. In the case of SF_{est}, the term may wrongly estimate the SF from some dominant scatterers under consideration, and may further neglect scattering from other sources (hereafter termed background scattering). In addition to the above, the shift-invariant convolution model may itself be inaccurate: there may be electrical noise on the image, or the PSF may vary spatially. Such variation will increase in the presence of nonlinear propagation. Last, the strength and concentration of scatterers may be high enough that multiple scattering becomes significant and the imaging process can no longer be described by a single convolution, even one that varies spatially [34].

DOI:10.15774/PPKE.ITK.2020.007

The next sections are as follows: first, an experimental setup is presented that allows a real ultrasound image to be generated and corresponding SFs to be estimated. The different methods for estimating the PSF are then presented. Last, the methods for generating the simulated images are introduced, including the evaluation of simulation accuracy and the role of different sources of error.

3.2.5 Generation of co-aligned US image and SF

General Experimental Setup

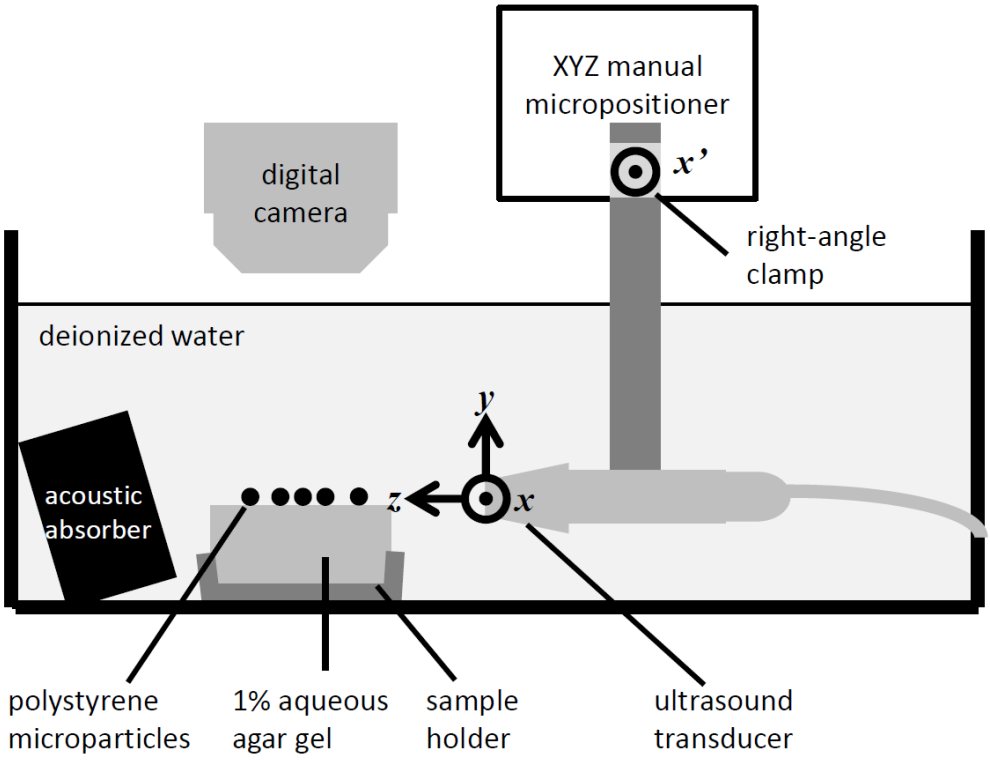

Figure 3.1: Schematic diagram of experimental setup. 49-μm-diameter polystyrene microspheres are resting on the surface of a 1% agar gel. The surface is in the imaging plane of a linear array that has been aligned with an XYZ micropositioner. Alignment was achieved by successive rotations around the x' axis, and movements in the y direction until the level of scattering from the polystyrene spheres, as observed on a live ultrasound B-mode image, was maximized. An acoustic absorber made of graphite-loaded PDMS and placed at an angle to the array surface is used to reduce multiple reflections. A digital camera in macro mode (1 cm focus) images the distribution of microspheres. [Th1]

Figure 3.1 shows a schematic of the experimental setup. A water tank with

DOI:10.15774/PPKE.ITK.2020.007

deionized water was used to acoustically couple the diagnostic ultrasound linear array (LA522E, Esaote, Genoa, Italy) with the 49-μm-diameter polystyrene microspheres (Chromosphere BK050, Thermo Scientific, Waltham, MA, USA) under investigation. The density and bulk modulus of polystyrene are 1.04 kg·m^{-3} and 6.2 MPa, compared with 1.00 kg·m^{-3} and 2.2 MPa for water [142], so that scattering is expected to be primarily from the compressibility contrast.

The microspheres had been suspended in deionized water and then placed on top of a flat 1% aqueous agar gel using a micropipette. The spatial distribution of the microspheres was imaged using a digital camera (SP-820UZ, Olympus, Tokyo, Japan) on super macro setting (1 cm focus). The agar concentration of 1% was chosen to make the gel acoustically transparent so that the microspheres appeared to float in the imaging plane of the ultrasound transducer. An acoustic absorber made of graphite-loaded polydimethylsiloxane gel and placed at an angle to the transducer surface was used to reduce multiple reflections in the ultrasound image.

Generation of ultrasound images

The 192-element linear array (LA522-E, Esaote, Genoa, Italy) had a 3 to 6 MHz response, 47 mm aperture, a 20 mm elevation focus depth, and was connected to an Ultrasound Advanced Open Platform (ULA-OP) Research US system (Microelectronics Systems Design Laboratory, University of Florence, Florence, Italy) [143]. Due to multiplexing, 64 elements (or 15.7 mm) of the aperture were active at any time. The imaging system allows recording of pre-beamformed data that can be used later to generate ultrasound images using arbitrary receive beamforming settings. The post-beamformed RF images thus obtained will be denoted by I_{real}, and the envelope-detected B-mode images by B_{real}. As in Fig. 3.1, the transverse, elevation, and axial directions in the image are denoted by x, y, and z, respectively.

To help ensure the validity of the shift-invariance assumption, the imager employed a uniform delay on transmit and dynamic receive beamforming. A reference ultrasound recording was taken before placement of the microparticles. As is typical for linear array imaging architectures, the contiguous, 64-element active subaperture stepped through the 192-element linear array in a consecutive manner. Due to the

DOI:10.15774/PPKE.ITK.2020.007

uniform-delay transmission, this allowed averaging of the pre-beamformed RF data up to 32 times, if needed.

The imaged region ranged from a depth z of 17.2 to 23.7 mm, close to the elevation focus depth. Because it is linearly proportional to the receive beamwidth, one parameter of interest was the receive F#, defined as the ratio of focal distance to the receive aperture. Unless otherwise stated, images were generated without dynamic receive apodization, using the maximum available aperture throughout (giving F# = 1.1—1.5). In the case of dynamic receive apodization, the maximum available aperture was used at the maximum imaging depth (giving F# = 1.5). For the calculation of beamforming delays and imaged depths z, the speed of sound was assumed to be 1482 m/s based on the temperature of the water (20 °C) [144].

Estimation of the scattering function

Three methods were used to estimate the SF, whose theory is described in Section 3.2.1. These methods corresponded to three image processing operations carried out on the macrophotograph obtained, as illustrated in Fig. 3.2.

1. $\mathrm{SF}_{\mathrm{threshold}}$: After inversion of the red channel, Otsu's method [145] was employed by the Matlab (The MathWorks Inc., Natick, MA, USA) built-in function graythresh to obtain a suitable threshold with which to separate the polystyrene spheres from the background. This threshold was used to convert the grayscale image into a binary image, which, after removal of objects with fewer than 10 pixels, was used to generate the estimate $\mathrm{SF}_{\mathrm{threshold}}$.

2. $\mathrm{SF}_{\mathrm{project}}$: In the knowledge that the circular cross-sections represent spheres, the distance transform was applied on $\mathrm{SF}_{\mathrm{threshold}}$ to produce $\mathrm{SF}_{\mathrm{project}}$. The distance transform converts circle functions into 2-D parabolic functions that represent the projection of a sphere onto a plane [36].

3. $\mathrm{SF}_{\mathrm{points}}$: Last, by finding local maxima in $\mathrm{SF}_{\mathrm{project}}$, the SF could be estimated as a set of point scatterers, yielding the estimate $\mathrm{SF}_{\mathrm{points}}$.

For reference, the 49-μm-diameter polystyrene spheres have a ka number of 0.49 when insonified at 4.7 MHz, the central frequency of the transducer.

DOI:10.15774/PPKE.ITK.2020.007

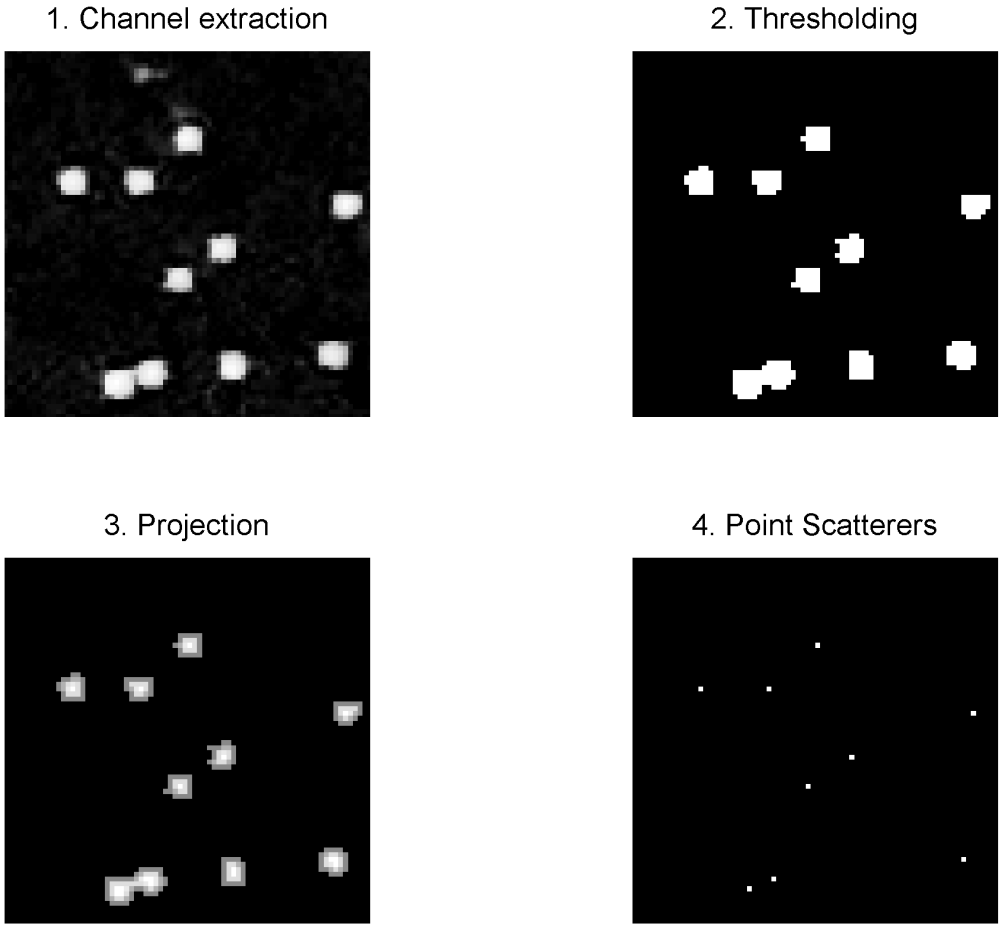

Figure 3.2: Steps in estimating the SF of the polystyrene scatterers from a macrophotograph. 1. Extraction of the red channel in the RGB photograph and inversion to produce a grayscale image. 2. Thresholding of the grayscale image. 3. Projection of sphere functions onto a 2-D image. 4. Reduction of scatterers to discrete points. The last three steps produce three corresponding estimates of SF, namely $\mathrm{SF}_{\mathrm{threshold}}$*,* $\mathrm{SF}_{\mathrm{project}}$*, and* $\mathrm{SF}_{\mathrm{points}}$*. [Th1]*

Registration of ultrasound image with SF

To align the ultrasound image with the SF, an estimate was necessary of the inter-pixel distance dx in the macrophotograph. Based on an initial estimate of 8300 nm from the dimensions of the agar gel (23.5 $\times$ 23.5 mm in area), the cross-correlation between $\mathrm{SF}_{\mathrm{threshold}}$ and the B-mode ultrasound image B_{real} was calculated for candidate values of dx in the range of 8000 nm to 8600 nm, in steps of 5 nm. To estimate dx, the value that maximized the spatial maximum of the cross-correlation function was chosen, whereas the location of the spatial maximum provided the spatial alignment between $\mathrm{SF}_{\mathrm{threshold}}$ and B_{real}.

DOI:10.15774/PPKE.ITK.2020.007

3.2.6 Estimation of the PSF

As described in Section 3.2.2, three possible techniques to estimate the PSF are the Field II-based, deconvolution-based, and data-based methods.

1. $\mathrm{PSF}_{\mathrm{FieldII}}$: The spatial impulse response of the transducer was estimated with Field II [128, 139] using transducer element parameters provided by the manufacturer. In addition to those already mentioned, these were as follows: 3 cycles at 4.7 MHz pulse-echo wavelet, 6 mm element height, 0.245 mm pitch, and negligible kerf. As mentioned in Section 3.2.5, imaging was only carried out using 64 of the 192 elements at any one time; this setting was also applied in the Field II simulations..

2. $\mathrm{PSF}_{\mathrm{deconvolution}}$: The US image I_{real} was deconvolved with $\mathrm{SF}_{\mathrm{threshold}}$ in the Fourier domain using Eq. (3.1).

3. $\mathrm{PSF}_{\mathrm{data}}$: The aligned ultrasound image I_{real} and scatterer function estimate $\mathrm{SF}_{\mathrm{threshold}}$ (as described in Section 3.2.5) was used to select a region of I_{real} that contained only one scatterer.

For all three methods, the PSF estimates were cropped around a window 0.84 mm long by 2.84 mm wide, which was chosen to encompass the main lobe of the PSF. This was particularly significant in the case of $\mathrm{PSF}_{\mathrm{deconvolution}}$, where without cropping, the simulation would have simply returned the real ultrasound image [Eqs. (2.60) and (3.1)]. Such an estimate, however, would have been an artificial means of achieving a perfect simulation, whereas cropping the estimate around the main lobe was expected to yield a physically realistic estimate. However, it was also feared that without smooth windowing, the sudden drop in amplitude at the edges of the PSF would artificially introduce high frequency components into the PSF estimate and thus the simulated ultrasound image. Therefore, the effect of multiplying the PSF estimates by a 0.84×2.84 mm Hanning window was also investigated. From here on, the use of the rectangular or Hanning windows will be indicated by the subscripts rect and Hanning, respectively.

DOI:10.15774/PPKE.ITK.2020.007

To compare the similarity between the different PSF estimates, the coefficient of determination R^2 was calculated between a reference estimate ($\mathrm{PSF}_{\mathrm{FieldII,rect}}$) and all the other estimates. This calculation was repeated for the envelopes of the PSF estimates to show the extent to which it was the phase or the amplitude of the estimates that differed.

A short summary of the PSF and SF estimates can be found in Table 3.1.

Table 3.1: Short summary of the PSF and SF estimates.

PSF estimates		SF estimates	
$\mathrm{PSF}_{\mathrm{FieldII}}$	estimated using Field II	$\mathrm{SF}_{\mathrm{threshold}}$	thresholding of the grayscale image
$\mathrm{PSF}_{\mathrm{deconvolution}}$	I_{real} deconvolved with $\mathrm{SF}_{\mathrm{threshold}}$	$\mathrm{SF}_{\mathrm{project}}$	projection of sphere functions from $\mathrm{SF}_{\mathrm{threshold}}$
$\mathrm{PSF}_{\mathrm{data}}$	one scatterer from I_{real}	$\mathrm{SF}_{\mathrm{points}}$	discrete points from $\mathrm{SF}_{\mathrm{project}}$ based on local maxima
$\mathrm{PSF}_{\{.\},\mathrm{rect}}$	rectangular-windowed PSF		
$\mathrm{PSF}_{\{.\},\mathrm{Hanning}}$	Hanning-windowed PSF		

3.2.7 Convolution-based ultrasound simulations

Simulation methods

Simulations were conducted in Matlab. A 2-D convolution was performed between $\mathrm{SF}_{\mathrm{est}}$ and $\mathrm{PSF}_{\mathrm{est}}$ using a third-party FFT-based implementation [146], yielding a simulated RF image I_{sim}. The magnitude of the analytic signal gave the envelope-detected image B_{sim}.

Before any further simulations were carried out, a method was sought to validate R^2 (see Section 3.2.3) as a measure of agreement between simulations and experiment. For this, a simulated image was generated using $\mathrm{SF}_{\mathrm{est}} = \mathrm{SF}_{\mathrm{threshold}}$, and $\mathrm{PSF}_{\mathrm{est}} = \mathrm{PSF}_{\mathrm{FieldII,rect}}$. The real and simulated ultrasound images were circularly shifted with respect to each other in units of pixels, whereupon R^2 was calculated for all possible shifts. R^2 was calculated both between the RF images I_{real} and I_{sim}, and between the envelope (B-mode) images B_{real} and B_{sim}, yielding the measures R_I^2 and R_B^2, respectively (see Section 3.2.3).

DOI:10.15774/PPKE.ITK.2020.007

Evaluation of Simulation Accuracy for Different Estimates of SF and PSF

In an attempt to find the best possible simulation accuracy, R_I^2 and R_B^2 were calculated for various estimates $\mathrm{SF}_{\mathrm{est}}$ and $\mathrm{PSF}_{\mathrm{est}}$. First, simulations were carried out using $\mathrm{SF}_{\mathrm{threshold}}$, $\mathrm{SF}_{\mathrm{project}}$, and $\mathrm{SF}_{\mathrm{points}}$ (see Section 3.2.5) and $\mathrm{PSF}_{\mathrm{FieldII,rect}}$ as a reference estimate of the PSF. Then, simulations were carried out using the three different estimates of the PSF, with $\mathrm{SF}_{\mathrm{threshold}}$ used as the estimate of the SF.

Identification of sources of error

A list of possible sources of simulation error have been provided in Section 3.2.4, whose relative roles were assessed by the variation in R_I^2 and R_B^2 for different simulation and imaging settings. The variation with different estimates of $\mathrm{SF}_{\mathrm{est}}$ and $\mathrm{PSF}_{\mathrm{est}}$ has already been addressed previously in Section 3.2.7. The role of background scattering in the simulation error was assessed by subtracting a reference ultrasound image (taken before scatterer displacement) from the real ultrasound image, whereas that of electrical noise was assessed by temporal averaging of the pre-beamformed data 32 times (see Section 3.2.5).

The use of uniform transmit delays was expected to reduce the shift-variance of the PSF; the effect of using dynamic receive apodization to further minimize shift-variance was also investigated. In addition to the issue of shift variance, the level of error may vary with the concentration of scatterers, either due to the varying level of signal, or due to the introduction of multiple scattering from the polystyrene spheres. These effects were investigated by calculating spatial maps of R_I^2 and R_B^2 over 40 $\times$ 40 pixel windows (corresponding to 333 $\times$ 333 nm) so that the spatial variation of simulation accuracy can be observed, including for regions of low and high scatterer concentration. Finally, the role of nonlinear propagation was assessed by calculating the ratio of second harmonic (9.4 MHz) to fundamental (4.7 MHz) signal amplitude for the imaged region of interest.

DOI:10.15774/PPKE.ITK.2020.007

3.3 Results and discussion

3.3.1 Generation and registration of ultrasound image with scattering function

Figure 3.3 shows the estimated scattering function $\mathrm{SF}_{\mathrm{threshold}}$ aligned with the real ultrasound image. During the process of alignment (see Section 3.2.5), the distance dx between the macrophotograph pixels was estimated as 8325 nm. The other estimates $\mathrm{SF}_{\mathrm{project}}$, $\mathrm{SF}_{\mathrm{points}}$ are not shown due to lack of visibility, but differences between them at a higher zoom level can be observed in Fig. 3.2.

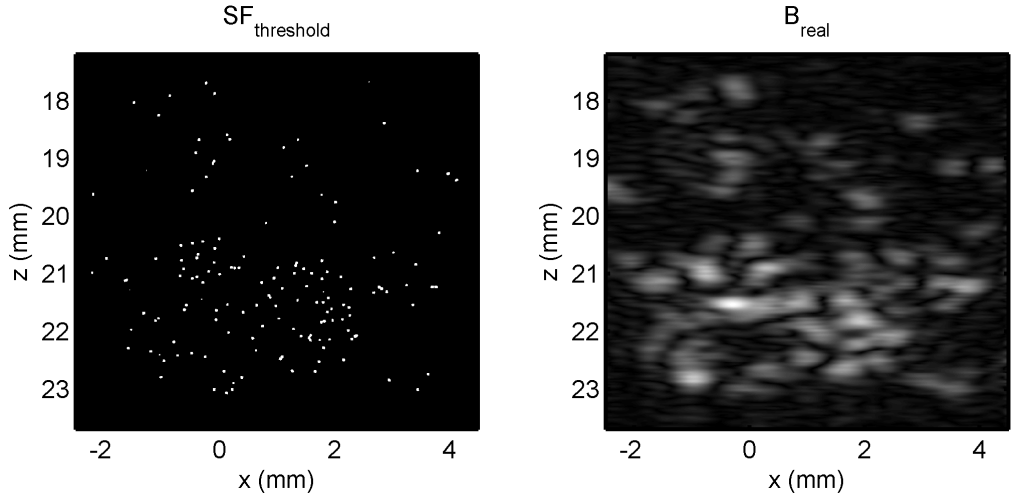

Figure 3.3: Alignment of estimated SF with experimentally obtained B-mode image. (left) Estimated scattering function $\mathrm{SF}_{\mathrm{threshold}}$ *based on thresholded macrophotograph data (Fig. 3.2). (right) Corresponding experimentally obtained B-mode image* B_{real}*. [Th1]*

3.3.2 Estimation of PSF

Figure 3.4 shows the various estimates of PSF (see Section 3.2.6), with values of R_I^2 and R_B^2 showing their variation in phase and amplitude compared with the reference $\mathrm{PSF}_{\mathrm{FieldII,rect}}$. The values of R_I^2 suggest significant differences between the reference PSF and other PSFs. On the other hand, there is less variation between the envelopes, implying that the estimates vary less in amplitude than in phase. Hanning windowing increases the similarity with the reference $\mathrm{PSF}_{\mathrm{FieldII,rect}}$, which means that there is considerable variation between the PSF estimates near the edges

DOI:10.15774/PPKE.ITK.2020.007

that is thereby suppressed. It is hypothesized that this variation is mostly from noise; the validity of this hypothesis will be evaluated in the next subsection, where the accuracy of simulations will be considered.

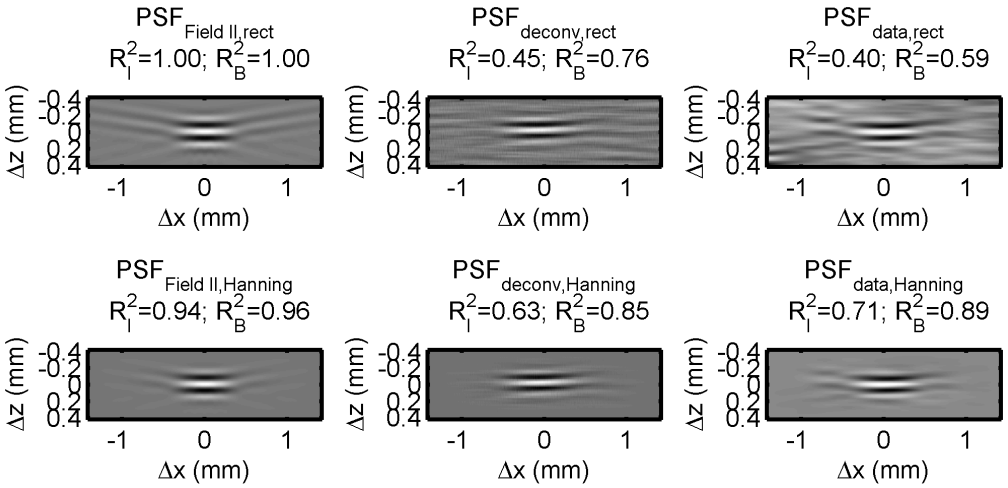

Figure 3.4: Comparison of different estimates of the PSF of the imaging system. As described in Section 3.2.6, three methods were tested: the Field II-based method (left column), the deconvolution method (center column), and the data-based method (right column). The PSFs were multiplied by a rectangular window (top row) and a Hanning window (bottom row). Using $\mathrm{PSF}_{\mathrm{FieldII,rect}}$ *as a reference, values of* R^2 *were computed for the PSFs (*R_I^2*) as well as for their envelopes (*R_B^2*). [Th1]*

3.3.3 Convolution-based ultrasound simulations

Figure 3.5 shows the variation of R_I^2 and R_B^2 as the real and simulated ultrasound images are spatially shifted with respect to each other. Both R_I^2 and R_B^2 drop to negligible levels for spatial shifts exceeding the dimensions of the PSF (Fig. 3.4), demonstrating their validity as measures of similarity between real and simulated images. Interestingly, R_B^2 is less spatially sensitive and shows higher values of agreement than R_I^2. These differences are deemed to be due to the sensitivity of R_I^2 to phase mismatch, which is demonstrated by the oscillatory variation of R_I^2 with axial shift Δz.

DOI:10.15774/PPKE.ITK.2020.007

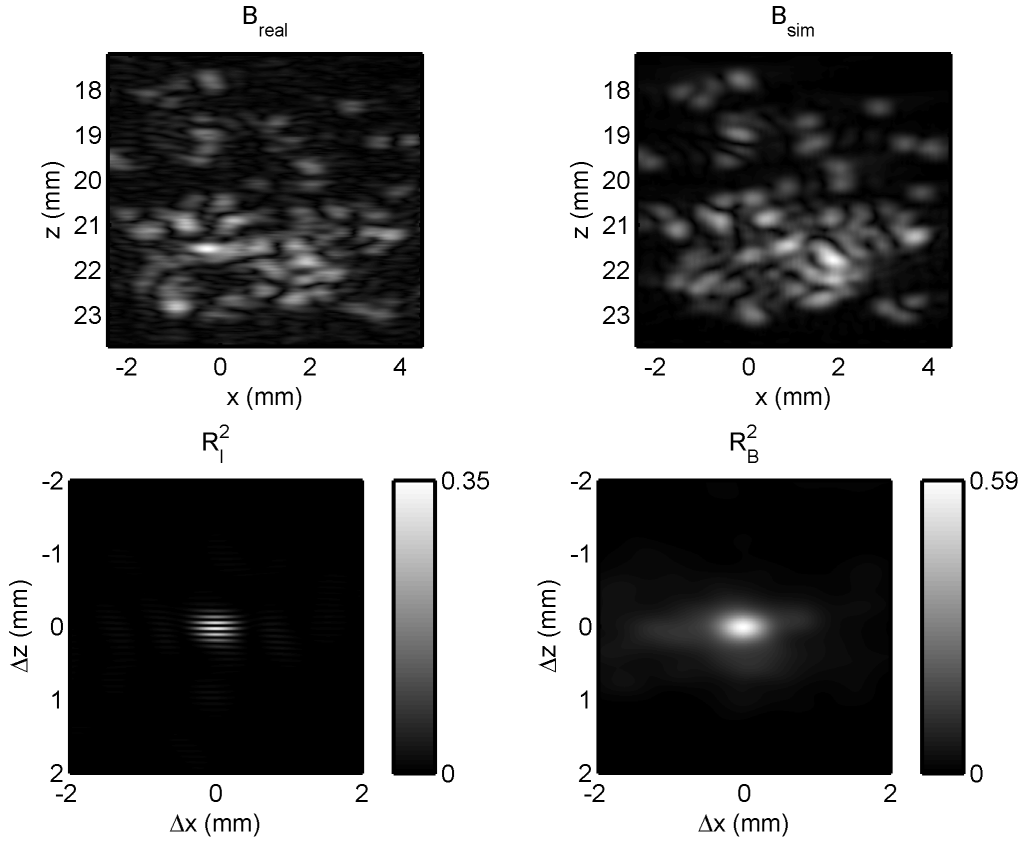

Figure 3.5: Spatial sensitivity of the coefficients of determination R^2 for the RF and envelope (B-mode) data between real and simulated ultrasound images. (top left) Real ultrasound image. (top right) Simulated ultrasound image computed using $\mathrm{SF}_{\mathrm{threshold}}$, $\mathrm{PSF}_{\mathrm{FieldII,rect}}$. *(bottom) Variation of R^2 for RF images (left) and envelope (B-mode) images (right) as a function of lateral and axial circular shifts Δx, Δz between real and simulated ultrasound images. Although the full range of circular shifts was computed, the figures only show a $[-2\ 2]$ mm box for greater clarity. Outside the box, values of R^2 are negligible. [Th1]*

Evaluation of simulation accuracy for different estimates of SF, PSF

Figure 3.6 compares the real ultrasound image and simulations obtained using three estimates of the scattering function: $\mathrm{SF}_{\mathrm{threshold}}$, $\mathrm{SF}_{\mathrm{project}}$, and $\mathrm{SF}_{\mathrm{points}}$. $R_I^2 = 0.36$ for all three simulations, whereas R_B^2 is in the range of 0.59 to 0.61, showing that there is little difference in simulation accuracy. This is likely because the polystyrene spheres are close to being point scatterers (ka = 0.49).

In Fig. 3.7, the real ultrasound image is compared with simulations obtained using the six different PSF estimates shown in Fig. 3.4. As expected, Hanning windowing the PSF estimate will suppress high-frequency components at the edges

DOI:10.15774/PPKE.ITK.2020.007

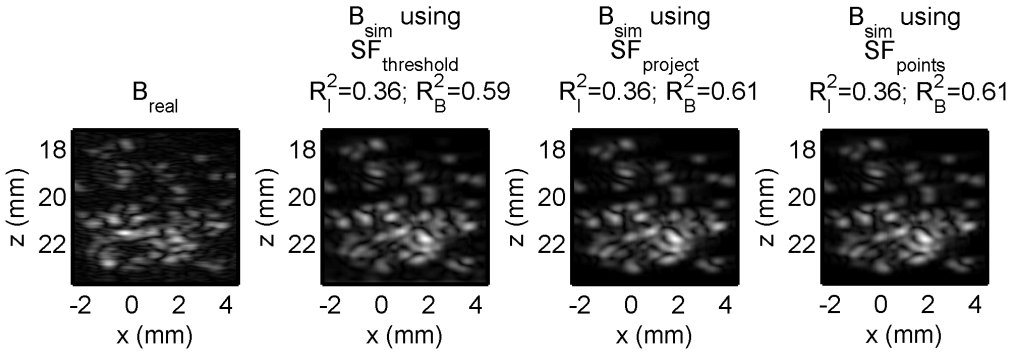

Figure 3.6: Comparison between the real ultrasound (left) and simulated ultrasound images computed using three different estimates of the SF (Fig. 3.2). All three simulations used $\mathrm{PSF}_{\mathrm{FieldII,rect}}$ *(Fig. 3.4). [Th1]*

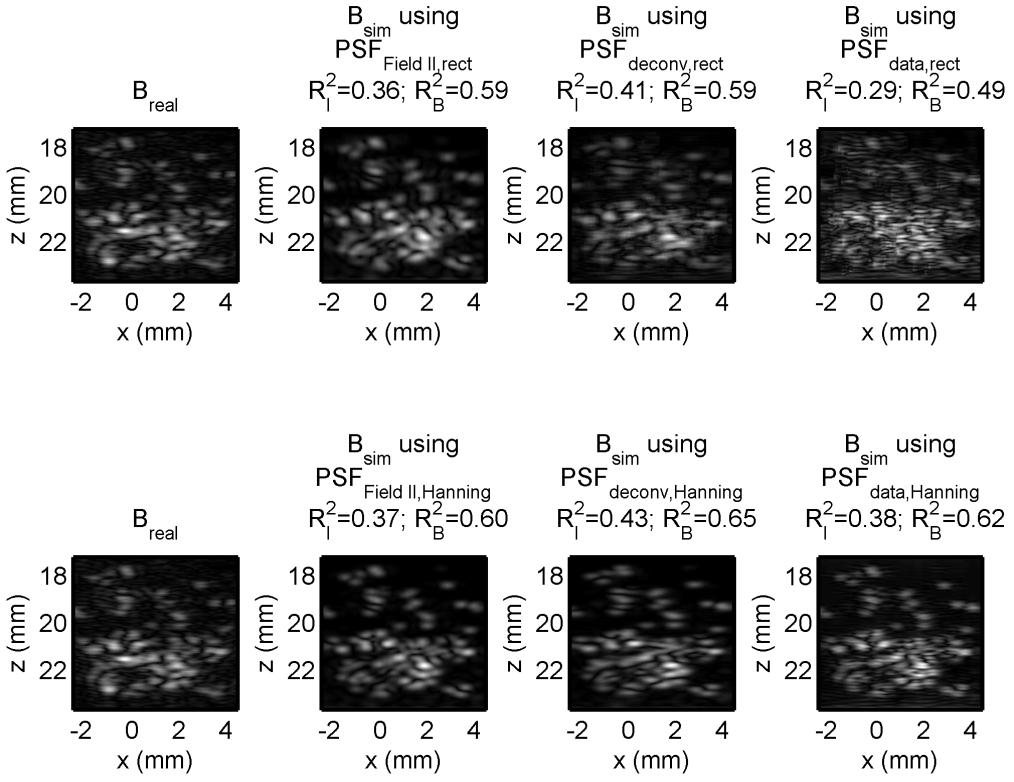

Figure 3.7: Comparison between the real ultrasound image (first column) and simulated ultrasound images computed using six different estimates of the PSF (Fig. 3.4). For all simulations, the SF was estimated by $\mathrm{SF}_{\mathrm{threshold}}$ *(Fig. 3.2). [Th1]*

(see Section 3.2.6) and suppress noise near the edges (see previously in Section 3.3.3), thereby yielding better simulation results. Compared with the reference estimate $\mathrm{PSF}_{\mathrm{FieldII,rect}}$, the best estimate $\mathrm{PSF}_{\mathrm{deconv,Hanning}}$ increases R_I^2 from 0.36 to 0.43 and

DOI:10.15774/PPKE.ITK.2020.007

R_B^2 from 0.59 to 0.65.

Identification of error sources

From previously in Section 3.3.3, it could be seen that there was little difference in simulations when using different estimates of SF and a somewhat greater difference (a change in R_I^2 of up to 0.07) with different estimates of the PSF. However, even with the best estimates, only 43% of the signal variation in the real ultrasound image could be explained by the simulation, or 65% in the case of envelope-detected images. Therefore, in addition to errors in estimating the SF and PSF, other sources of error are also likely to be significant.

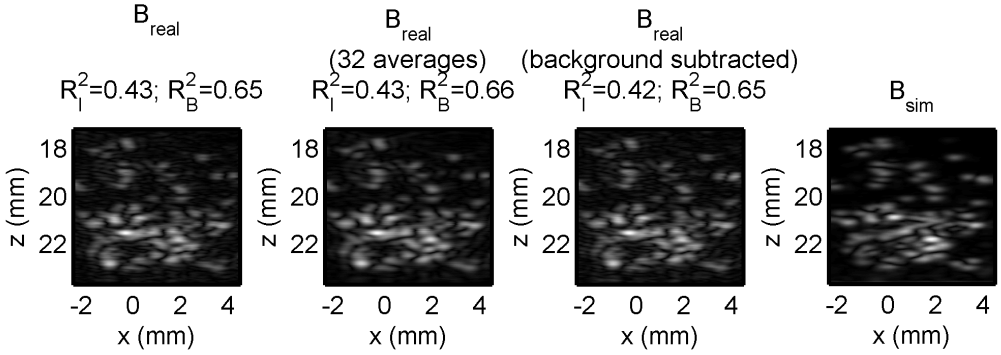

Figure 3.8: Comparison between real ultrasound images and a simulated image. The real ultrasound images were generated without averaging or reference subtraction, with averaging of the pre-beamformed RF data to reduce electrical noise, and with subtraction of a reference signal before scatterer placement to reduce background scattering. The simulated image was generated using $\mathrm{SF}_{\mathrm{threshold}}$ *and* $\mathrm{PSF}_{\mathrm{deconv,Hanning}}$ *(Fig. 3.7). [Th1]*

Figure 3.8 shows the effect of temporal averaging and of subtracting a background image without polystyrene scatterers. The resulting changes in R_I^2 and R_B^2 are at most 0.01, showing that the contributions of electrical noise and background scattering are minimal. Figure 3.9 shows the effect of introducing dynamic receive apodization, which reduces shift variance in the PSF. Instead of an improvement, there is a slight decrease (0.02) in R_B^2, possibly due to a decrease in signal owing to a decrease in the receive aperture (see Section 3.2.5).

To help addressing the question of how shift variance and scatterer concentration affect simulation accuracy, Fig. 3.10 shows the spatial variation of R_I^2 and R_B^2.

DOI:10.15774/PPKE.ITK.2020.007

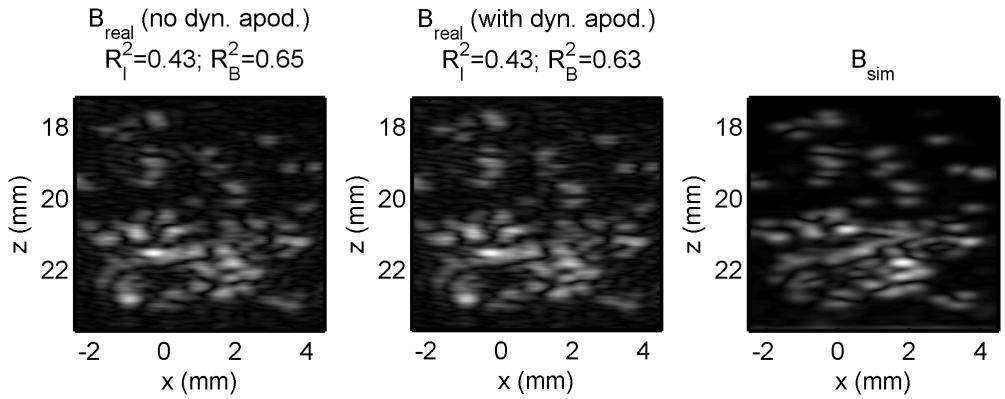

Figure 3.9: Comparison of two real ultrasound images, one generated without, the other with dynamic receive apodization, with a simulated ultrasound image. The simulated image was generated using $\mathrm{SF}_{\mathrm{threshold}}$ *and* $\mathrm{PSF}_{\mathrm{deconv,Hanning}}$ *(Fig. 3.7). [Th1]*

Both R_I^2 and R_B^2 display higher values in regions with scatterers, up to 0.98 and

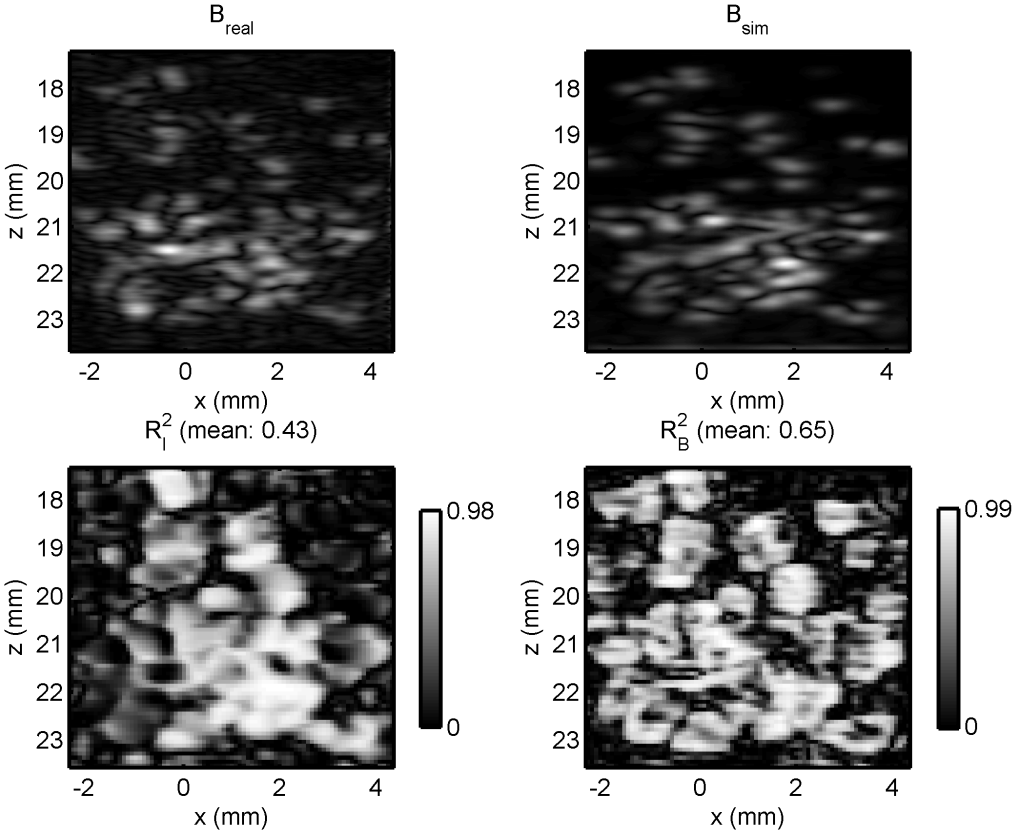

Figure 3.10: Estimation of the spatial variation of the coefficients of determination R^2 *for both the RF images* I *and B-mode images* B*. (top left) Real ultrasound image. (top right) Simulated ultrasound image, generated using* $\mathrm{SF}_{\mathrm{threshold}}$ *and* $\mathrm{PSF}_{\mathrm{deconv,Hanning}}$ *(Fig. 3.7). (bottom left) Spatial map of* R^2 *between* I_{real} *and* I_{sim}*. (bottom right) Spatial map of* R^2 *between* B_{real} *and* B_{sim}*. [Th1]*

DOI:10.15774/PPKE.ITK.2020.007

0.99, respectively. Shift variance in the axial direction is not noticeable. Nonlinear propagation could have caused significant spatial variation of the PSF in the axial direction by the introduction of harmonics; however, the ratio of the second harmonic (9.4 MHz) to fundamental (4.7 MHz) signal amplitude was measured as 0.05. It is likely that the reason behind this relatively low ratio compared with typical diagnostic imaging scenarios [147, 148] is that uniform delay transmission was employed, which reduced peak pressure amplitudes.

In contrast to the lack of axial shift variance, the lower values of R_I^2 around the left and right edges (not present in the R_B^2 map) are indicative of phase decorrelation. Stretching the estimate of SF in the transverse direction did not improve results. It is possible, however, that misalignment between the SF and real ultrasound image was more complex, such as a rotation error. Because R_I^2 has been shown to be more sensitive to alignment errors than R_B^2, this would explain why R_B^2 retains higher values at the left and right edges.

The results of Fig. 3.10 show that a higher concentration of scatterers did not reduce simulation accuracy, which implies that in the current experiment, multiple scattering from polystyrene scatterers was not a significant phenomenon. On the contrary, low scatterer concentrations reduced the level of signal and therefore simulation accuracy. It is not clear, however, what source of noise is chiefly responsible for the reduction in simulation accuracy. Figure 3.8 shows that the roles of electrical noise and background scattering are not dominant. One hypothesis is that reflections from the acoustic absorber behind the phantom (Fig. 3.1) may have caused additional scattering from the polystyrene particles. The other is that the air content of the water may have caused microbubbles to appear over time, which could not be adequately subtracted using the reference image. One such microbubble may be at x = −0.5 mm, z = 21.5 mm (Fig. 3.10).

3.4 Conclusions

The present chapter has shown how a shift-invariant model of ultrasound image formation is able to predict ultrasound images of microparticles in the imaging plane

DOI:10.15774/PPKE.ITK.2020.007

of a linear array. Using the coefficient of determination metric allowed the calculation of the proportion of variance in the real ultrasound image that could be explained by the simulated image. The coefficient of determination was calculated both for the RF ultrasound image and the envelope-detected (B-mode image), denoted by R_I^2 and R_B^2, respectively. After the SF estimate was aligned with the ultrasound image, the variation of both measures was calculated as the real and simulated images were circularly shifted with respect to each other. R_I^2 and R_B^2 exhibited clear maxima at the correct alignment, diminishing to negligible levels when the misalignment was greater than the dimensions of the PSF. In addition, R_I^2 was shown to be more sensitive to misalignment errors than R_B^2 due to the former being sensitive to phase differences. These results underlined the validity of the coefficient of determination measure, and the differing interpretations of R_I^2 and R_B^2.

In the next step, various estimates of SF, PSF were tested to see which yielded the best simulation. All three estimates of SF gave similar results. In the case of the PSF, deconvolution of the real image with the estimated SF, followed by Hanning windowing ($\mathrm{PSF}_{\mathrm{deconv,Hanning}}$) gave the best results, yielding $R_I^2 = 0.43$, $R_B^2 = 0.65$.

In the last step, the source of simulation error was explored. Regions of scatterers showed high values of R_I^2 and R_B^2 (up to 0.98 and 0.99, respectively), whereas regions without scatterers were dominated by noise. Because the reduction of electrical noise and background (nonpolystyrene) scattering from agar only changed values of R_I^2 and R_B^2 by up to 0.01, it is possible that other sources of error were more significant. For example, scattering of the polystyrene particles from multiple reflections or scattering from microbubbles could have also contributed to simulation errors. From the above observations, it is expected that by increasing the concentration of imaged scatterers or by more careful experimental design, higher overall values of R_I^2 and R_B^2 can be obtained.

To the best knowledge of the author, this work presents the first example of a medium with an estimated SF being imaged with ultrasound, allowing for the validation of the shift-invariant convolution model of ultrasound image formation. The above results show that, at least for the experimental setup used in the current work, the shift-invariant convolution model describes most of the variation in

DOI:10.15774/PPKE.ITK.2020.007

a B-mode image; however, care should be taken to reduce other sources of scattering such as multiple reflections or microbubbles. Planned extensions of this work include validating the model using other scatterer arrangements such as 3-D distributions of spherical scatterers, live biological cells, and eventually tissue. In addition to validating the shift-invariant convolution model, the experimental setup herein provided allows a host of experiments to be conducted where knowledge of the SF is required. For instance, methods to estimate any structural descriptor may be tested in an experimental setting, including the location of individual scatterers using superresolution imaging [149].

DOI:10.15774/PPKE.ITK.2020.007

Chapter 4

Resolution enhancement of high-resolution B-mode images using axial processing

4.1 Introduction

Enhancement of image resolution in ultrasound images is key to help clinicians find early indicators of pathological lesions and has therefore long been of interest [150–152]. Image resolution enhancement relies on deconvolving the point spread function (PSF) of the imaging system out of the ultrasound image prior to envelope detection and other post-processing steps. Unfortunately, in most cases the PSF is spatially variant, complicating its estimation and subsequent use in deconvolution.

The aim of this chapter is to present how to simultaneously increase both the axial and lateral resolution of B-mode ultrasound images (acquired by a SAM setup) by using axial deconvolution only. The current piece of work is driven by the realization that the PSF at a given coordinate can be decomposed into spatially invariant and variant components, as detailed in Section 2.3.6 (Eq. (2.92)). If the bandwidth is increased more at the upper cut-off frequency, then the effective central frequency will also be increased, resulting in a better lateral resolution, improving US image resolution in both directions.

DOI:10.15774/PPKE.ITK.2020.007

4.2 Methods

To show that the axial deconvolution can change the lateral resolution, 4 different B-mode images were considered, 2 simulations and 2 measurements. Concerning the simulations, a sparsely and a densely populated scatterer area were simulated in turn; as for the measurements, a B-mode image of an agar-graphite phantom and a skin examination will be shown. Conventionally, the axial direction is indicated by z, while the lateral direction is annotated by x.

4.2.1 Simulations

The number of scatterers was set to 5 and 300 for the sparse and the dense areas, respectively. The characteristics of the simulated transducer match the single element transducer used for the agar-graphite measurement (see Section 4.2.2). The sampling frequency was set to 250 MHz, and the lateral step size to 1 μm.

4.2.2 Agar-graphite phantom

The measurement was carried out using a previously described SAM system [Au4]. A single-element transducer was used with a nominal center frequency of 35 MHz (PVDF Immersion Transducer PI35-2-R1.0, Olympus, Tokyo, Japan). The sampling frequency was set to 250 MHz and the lateral step size to 50 μm. The custom-made agar-graphite phantom contained deionized water, 3% agar and 4% graphite.

4.2.3 Skin examination

A single-element spherically focused transducer was used with a nominal center frequency of 20 MHz (Olympus V317, Tokyo, Japan, diameter: 6.3 mm; F#: 3). The sampling frequency was set to 80 MHz and the gain to 60 dB. A portable ultrasound system specifically made for skin diagnosis was used. B-mode ultrasound images were made of human skin. The image was acquired as part of an ethically approved study at the Department of Dermatology, Venereology and Dermatooncology, Semmelweis Medical University (Budapest, Hungary).

DOI:10.15774/PPKE.ITK.2020.007

4.2.4 Deconvolution

Two different methods were used for the deconvolution part, both implemented in MATLAB (Natick, WA). Firstly, the classical Wiener filter approach was applied with an assumed NSR level of 0.2 based on the real measurements. Secondly, a custom Fourier domain method (frequency-weighted axial filtering, RAMP) was applied, where the signal energy was boosted with a gradually increasing function at those (higher) frequencies, where the ultrasound transducer has a weaker response. Both of the methods were used along every A-line separately.

In the case of the simulations, the ZP impulse response was used for the deconvolution process. For the agar-graphite phantom and the clinical measurements the impulse responses were estimated as follows: the ZP mean Fourier spectrum along the axial direction was calculated and transformed into temporal domain, making this the shift-invariant component — assumed to be $v_{pe}(\mathbf{r})$ — of the pulse-echo response we sought to deconvolve out of the image.

4.2.5 Scaled RAMP filtering in the frequency domain

Another approach for resolution enhancement is to perform frequency-weighted axial filtering, where the Fourier spectrum of the reflected signal is multiplied with a gain function on higher frequencies. The above mentioned linear RAMP function $R(f)$ in our case was set to start with a scale of 1 (s_{start}) and end with a scale of 3 (s_{end}) over the frequency interval to be modified ($[f_{start}\ f_{end}]$). In mathematical terms:

$$R(f) = \begin{cases} \frac{f - f_{\text{start}}}{f_{\text{end}} - f_{\text{start}}} \cdot (s_{\text{end}} - s_{\text{start}}) + s_{\text{start}} & f \in [f_{\text{start}}\ f_{\text{end}}] \\ 1 & \text{otherwise} \end{cases} . \tag{4.1}$$

The multiplication with the gain function $R(f)$ in the Fourier spectrum was used on every A-line independently. The frequency intervals were defined empirically and are marked in a square bracket in the figures.

DOI:10.15774/PPKE.ITK.2020.007

4.2.6 Resolution estimation

For quantitative evaluation, the FWHM of the mean AC curves (both axial and lateral directions) of the envelope-detected B-mode US images were calculated to estimate the observed resolution. During the latter case the envelope of the mean AC curve was used. The area of the resolution cell was also approximated as the area of an ellipse.

4.3 Results and discussion

4.3.1 Simulations

Figures 4.1 and 4.2 show that Wiener deconvolution improved the axial resolution, however, resulted in a worse lateral resolution, while RAMP was able to increase resolution in both directions (1.5 times and 1.8 times better than original in axial and lateral directions, respectively). RAMP and Wiener methods showed

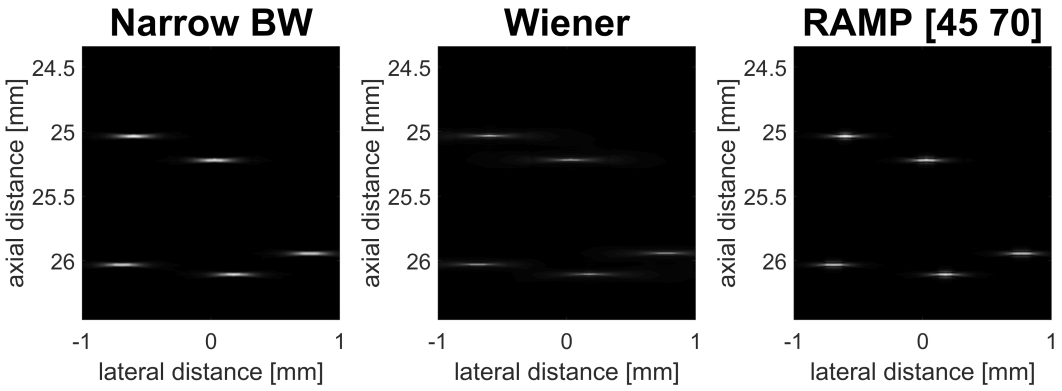

Figure 4.1: Resolution enhancement on a sparsely populated area (simulation) using deconvolution. Note the lateral smearing as a result of the Wiener filter. [Th2].

similar performance axially. Figures 4.3 and 4.4 show the AC curves of the sparse and dense areas, respectively. As it can be seen, while the axial performance of the RAMP and Wiener methods are similar, the lateral performance of the former greatly improved the lateral resolution too. In contrast, the Wiener deconvolution sacrificed the lateral resolution in order to improve the axial, resulting in lateral smearing (see Figs. 4.1 and 4.2).

DOI:10.15774/PPKE.ITK.2020.007

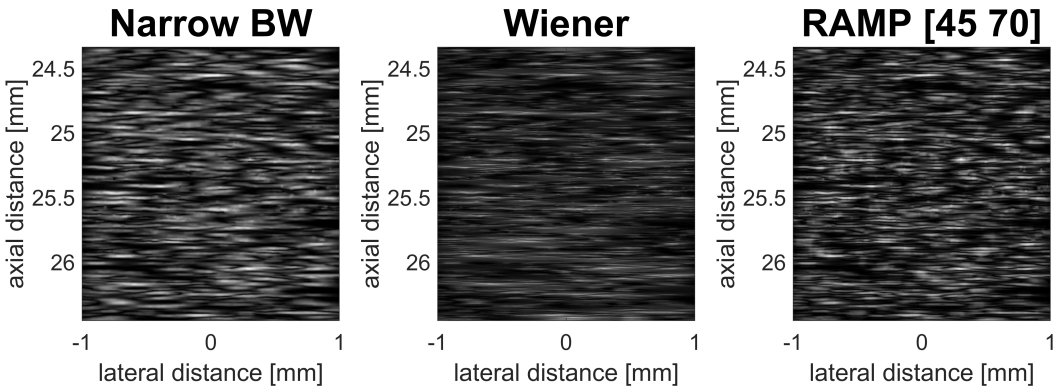

Figure 4.2: Resolution enhancement on a densely populated area (simulation) using deconvolution. Note the lateral smearing as a result of the Wiener filter. [Th2]

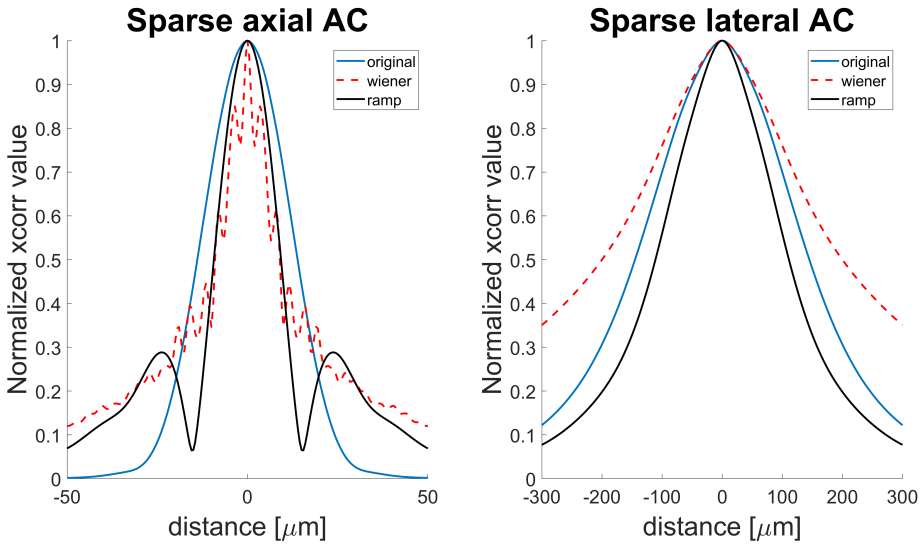

Figure 4.3: Axial and lateral mean AC curves of the envelope images (sparse area). In the lateral direction the envelope of the AC curve is shown. While RAMP method shows improvement in both directions, the Wiener method indicates improved axial resolution at the expense of a decrease in lateral resolution.[Th2]

4.3.2 Agar-graphite phantom

As Figs. 4.5 and 4.6 show, both methods could increase the resolution in both lateral and axial directions, however, resulted in amplified noise, which is much more noticeable using the Wiener method and could influence the resolution estimation.

DOI:10.15774/PPKE.ITK.2020.007

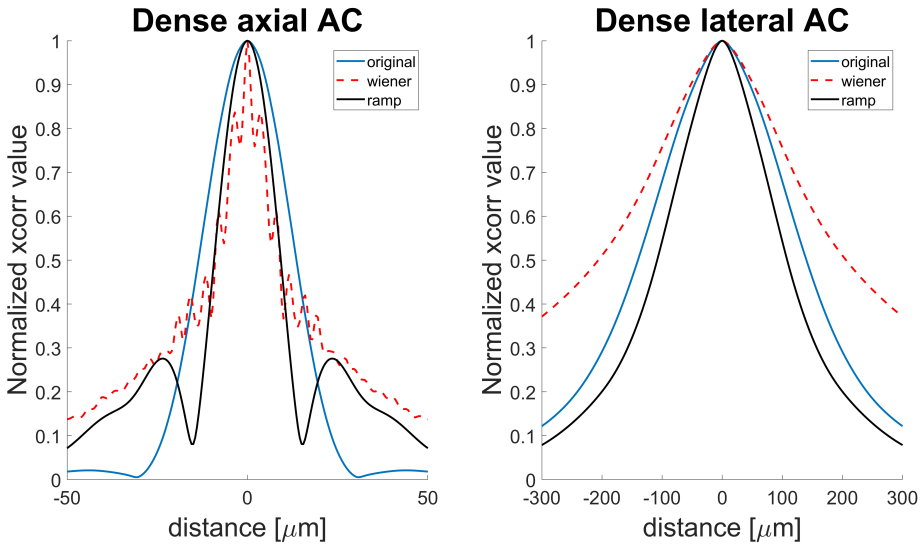

Figure 4.4: Axial and lateral mean AC curves of the envelope images (dense area). The results are similar to the sparsely populated area. In the lateral direction the envelope of the AC curve is shown. While RAMP method shows improvement in both directions, the Wiener method indicates improved axial resolution at the expense of a decrease in lateral resolution. [Th2]

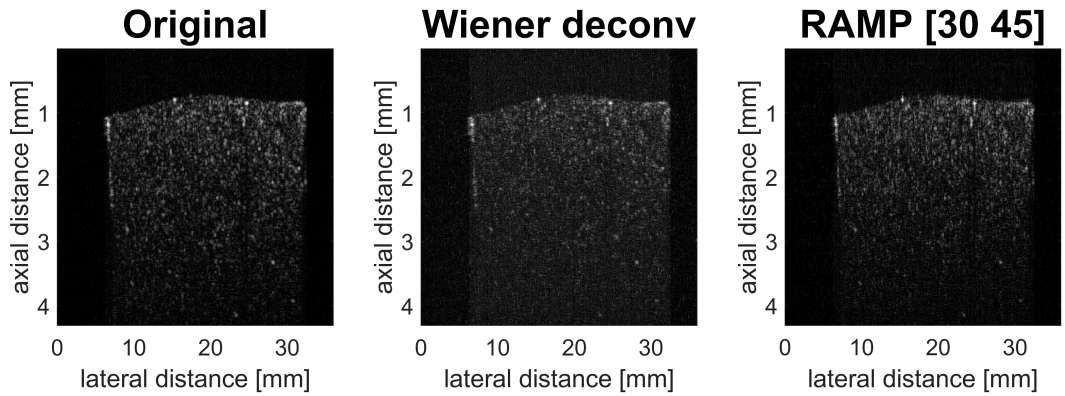

Figure 4.5: Resolution enhancement on an agar-graphite phantom B-mode US image using axial deconvolution. [Th2]

4.3.3 Skin examination

Figures 4.7 and 4.8 show that the Wiener method yields a better resolution in both directions, while the RAMP method greatly improved the lateral resolution at the cost of a slight deterioration of the axial resolution.

DOI:10.15774/PPKE.ITK.2020.007

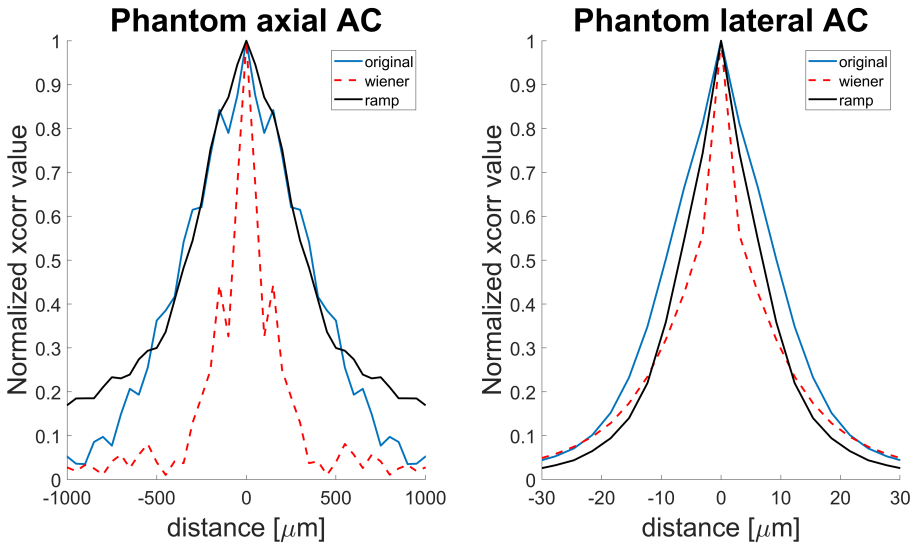

Figure 4.6: Axial and lateral mean AC curves of the envelope images (tissue-mimicking phantom). In the lateral direction the envelope of the AC curve is shown. The Wiener method shows a great improvement in axial direction, however, care needs to be taken to properly evaluate it as a real improvement due to the increased noise level. The RAMP method shows improvement in both directions. [Th2]

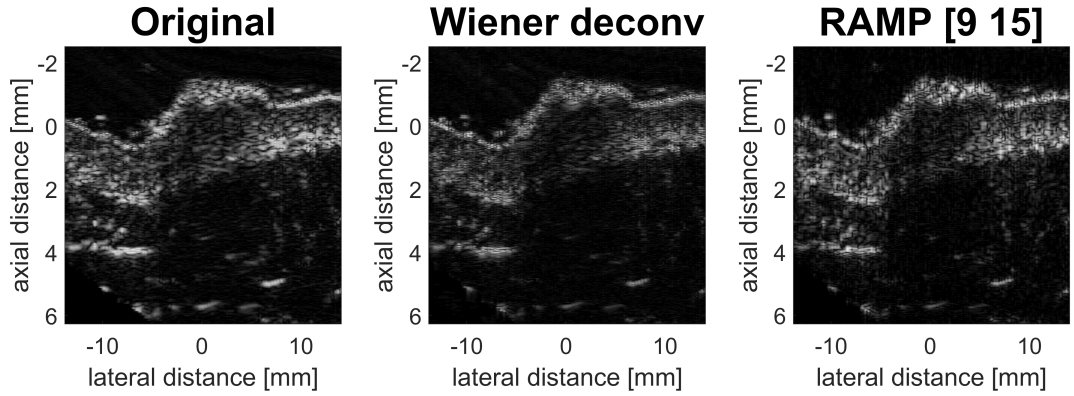

Figure 4.7: Resolution enhancement of a skin (nevus) B-mode image using deconvolution. [Th2]

4.3.4 Resolution estimation

Table 4.1 summarizes the resolution results obtained using the 4 US images for both deconvolution methods and compared with the original. Generally, it can be stated that both deconvolution approaches are able to improve lateral as well as axial resolution, though the RAMP filter tends to provide a better lateral resolution at the cost of a smaller improvement in axial resolution. The Wiener method

DOI:10.15774/PPKE.ITK.2020.007

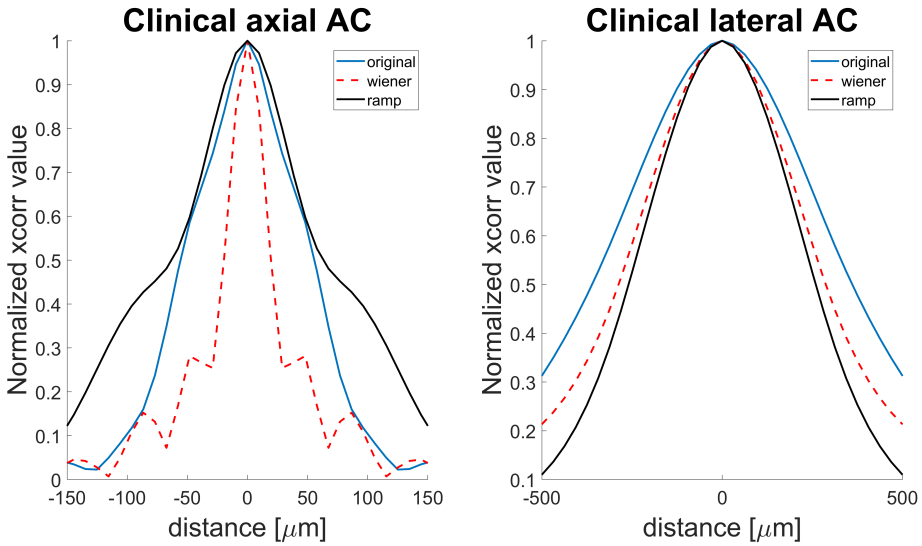

Figure 4.8: Axial and lateral mean AC curves of the envelope images (clinical skin image). The figure shows that the Wiener method yields a better resolution in both directions, while the RAMP method greatly improved the lateral resolution at the cost of a slight deterioration of the axial resolution. [Th2]

Table 4.1: FWHM values of the AC functions in μm (lateral x and axial z), and area of the resolution cell ($x \cdot z \cdot \pi$) in μm^2. It can be seen that the axial-lateral resolution cell (estimated as the area of an ellipse) always improved using the RAMP method. Table adapted from [Th2].

	orig		deconv		RAMP	
	x	z	x	z	x	z
	$x \cdot z \cdot \pi$		$x \cdot z \cdot \pi$		$x \cdot z \cdot \pi$	
sparse	290.0	27.8	399.8	18.0	222.1	18.7
	25327.5		22608.2		13047.9	
dense	280.4	27.2	412.1	18.0	216.4	18.6
	23960.6		23303.7		12645.0	
phantom	736.0	18.7	152.0	9.0	674.0	14.0
	43238.4		4297.7		29644.1	
skin	723.4	111.7	576.0	39.7	521.0	127.1
	253852.6		71839.4		208033.4	

shows a great improvement for real experiments; however, care needs to be taken to properly evaluate the results as a real improvement due to the increased noise level. In the case of the skin examination using the RAMP filter the axial resolution slightly decreased, yet the area of the resolution cell decreased as well, which is an acceptable trade-off taking the usually better axial resolution (compared to the

DOI:10.15774/PPKE.ITK.2020.007

lateral resolution) into account.

4.4 Conclusions

Using simulated and experimental data from two single element transducers (of 20, 35 MHz nominal frequencies), it was shown that axial deconvolution (both classical Wiener and a RAMP filter) can simultaneously improve resolution in both directions. The results demonstrate a framework for improving axial and lateral resolution for ultrasound images that is unaffected by depth-dependent effects and that can balance the need for axial and lateral resolution improvement based on their relative values. The results also highlight the need to set deconvolution parameters correctly. For cases when the deconvolution only improves lateral resolution at the expense of axial resolution, this may still result in a reasonable trade-off regarding the area of the resolution cell given the typically worse lateral resolution. Since both deconvolution methods introduce noise, care needs to be taken to keep it under control.

Future work can aim to optimize deconvolution parameters (NSR and RAMP characteristics) according to pre-defined axial and lateral resolution improvement criteria, or to even combine these methods in order to achieve a better performance.

DOI:10.15774/PPKE.ITK.2020.007

Chapter 5

Resolution enhancement of C-mode SAM images

5.1 Introduction

Ultrasound is a well-established diagnostic imaging modality that can image the whole depth of tissue at resolutions comparable to other diagnostic imaging modalities, such as PET, CT or MRI [58]. Its higher frequency counterpart, SAM, typically uses a single-element transducer of high central frequency that mechanically scans a volume of interest [153]. Due to the high frequencies used, SAM imaging is capable of producing high-resolution images [153–156], which is particularly useful for non-destructive testing [9–19] and examining biological tissues [155,157–167] or even single cells [15,168–174]. However, the increase in central frequency not only makes system manufacture more expensive but also decreases penetration depth. Use of a lower frequency allows for more rapid scanning of a tissue volume, since the higher penetration depth allows for thicker specimen slices. Moreover, scanning time of a single slice can also be reduced due to fewer required grid points. In general, SR techniques may be used to overcome this resolution limit. SR imaging either requires spatio-temporally discrete events [175] or post-processing of one [176, 177] or more images [178–180]. In increasing order of complexity, these post-processing techniques can be either deconvolution [19,181] [Th2], sparsity promoting minimization (*e.g.*, dictionary- [182] or TV-based [177]) or DL [183–186]. The more complex

DOI:10.15774/PPKE.ITK.2020.007

the method, the more computational power and running time are needed.

While ultrasound image resolution enhancement is the subject of interest in ongoing research [85–87,91,114,150–152,176,187–196] [Th1], there is currently limited research on the use of image resolution enhancement techniques in SAM imaging. Mohammadi and Mahloojifar [197] attempted to improve the axial resolution of SAM images using sparse signal representation by decomposing A-lines into overcomplete Gabor dictionaries. Basarab *et al.* [177] proposed a TV regularization-based method for single image SR with promising results. However, as the authors noted, there was a lack of ground truth with which they could compare their results.

The aim of this chapter is to compare two classical deconvolution-based and a DL-based image resolution enhancement method applied to 180-MHz centre-frequency SAM images, with SR performance measured by comparing with the corresponding 316-MHz data used as ground truth.

The structure of the chapter is as follows. First, the details of the experimental framework and data acquisition are described. Then, the image resolution enhancement methods are presented along with the metrics used for comparison. Then, the results are discussed, comparing the performance of the different SR techniques. Finally, conclusions are drawn based on the results.

5.2 Methods

In this section, the data collection for the experiment is first presented, followed by the details of the resolution enhancement methods. Finally, a description of the metrics used to measure resolution enhancement performance is given. All data processing steps were carried out using MATLAB (Mathworks, Natwick, MA, USA) unless otherwise stated.

DOI:10.15774/PPKE.ITK.2020.007

5.2.1 Experiment

Sample preparation

Five different samples of rat brain, pituitary gland and mouse brain sections were processed and imaged for the current experiment at well-identifiable anatomical landmarks (such as cerebral cortex, anterior pituitary lobe, or hippocampus) for accurate registration and contained vessels and cells of various densities. Perfusion-fixation was employed in all samples. The sections were mounted onto Superfrost UltraPlus slides and (due to a separate investigation) were stained either with Biot-Tomato Lectin (#LOG51; 1:100), CY3-streptavidine (1:1000), DAPI or Ni-Diaminobenzidine. Preliminary swelling and detachment tests were carried out in deionized water to ensure the sample maintained its conformation during the measurements. The thickness of the sections was around 30 μm.

Data acquisition

Measurements were carried out at Fraunhofer IBMT (Sulzbach, Germany) using custom-made spherically focussed single-element transducers (Fraunhofer IBMT) housed in a previously described SAM system (SASAM [170]).

Every recording was performed using two different transducers of central frequencies 180 and 316 MHz, whose characteristics can be seen in Table 5.1. The coupling medium was deionized water, held at a room temperature of 22 °C. Step sizes of 3 and 2 μm were used during imaging with the 180- and 316-MHz transducers, respectively. The focal plane of the transducers was set to the surface of the samples and the sampling frequency was 4 GHz.

Table 5.1: Properties of the transducers used during the experiment and PSF simulation.

Central frequency (MHz)	180	316
-6 dB bandwidth range (MHz)	132 – 230	242 – 420
Focal distance (μm)	555	425
Diameter of the curvature (μm)	482	319
Lateral resolution limit (μm)	9.6	6.4

DOI:10.15774/PPKE.ITK.2020.007

Due to the different step sizes, the 180-MHz image was interpolated to have the same spatial resolution as the image obtained at 316 MHz, using a common step size of 2 μm. Between the two measurements, the transducers needed to be manually changed; therefore, the images were co-registered as a post-processing step to match the scanned regions. The images were aligned using rigid transformation (only translation and rotation operations were allowed) with the help of manually selected control points. Additionally, for a more accurate registration, a cross-correlation-based alignment was performed between images at different frequencies to ensure one-to-one comparison.

5.2.2 Image resolution enhancement

Three different image resolution enhancement techniques were tested, as detailed in the following. First, two classical deconvolution techniques are presented, followed by the description of the DL method.

Classical deconvolution (Wiener and TV)

To be able to perform deconvolution the PSF needs to be estimated. To this end, a step-based estimation (SBE) was considered based on the actual data. The step response of the transducer was approximated by calculating the mean reflection from the glass slide over a 50 μm × 50 μm area. Our SBE of the impulse response was obtained by taking the derivative of this impulse response.

The SBE response was first multiplied by a modified Hanning window, which left the inner part of the signal unchanged, to reduce edge effects. Then, using the SBE response as an impulse response of a transducer, the PSF was simulated at the focal plane (z = 0) using the characteristics of the 180-MHz transducer (see Table 5.1) within an area of 50 μm × 50 μm using a step-size of 2 μm with the help of Field II [128, 139] in the x- and y-directions, respectively.

The Wiener deconvolution was performed independently slice by slice (as proposed by Basarab *et al.* [177]) using the built-in MATLAB deconvwnr() function, while the TV deconvolution was performed using a third-party function decontv() written by Chan *et al.* [198]. The SNR level for the Wiener deconvolution was set to

DOI:10.15774/PPKE.ITK.2020.007

10 dB, which was found to give the best results both qualitatively and quantitatively. Similarly, a regularization parameter value of 1 was used for the TV deconvolution based on empirical evidence. One common form of SAM imaging involves taking the integral of the absolute values of A-lines (from now on referred to as a C-scan SAM image, see Section 2.1.4); therefore, in this work, the 3-D volumetric data was turned into a 2-D image using the aforementioned procedure.

Deep learning

As mentioned in Section 2.3.5, DL relies on a training set, where the input images are paired with the corresponding output (reference) images in order to learn how to estimate the output based on the input. Ten SAM C-scan images were used in total, as described in the following. The initial training set included 4 full-size image pairs of 180 MHz and 316 MHz and had the following dimensions in the x- and y-directions, respectively [see Fig. 5.2 (from top to bottom) later in Section 5.3]: 642 μm × 848 μm, 800 μm × 796 μm, 798 μm × 798 μm, 698 μm × 698 μm, while the test set contained one full-size image pair and had an area of 420 μm × 700 μm (see Fig. 5.3 later in Section 5.3). Note that the difference in dimensions originates from the registration process between different frequencies (Section 5.2.1). The image with the smallest dimensions was selected as the test set to obtain the biggest training set possible. The test and training sets were carefully chosen to be disjunctive.

To create a properly sized training set, the full-sized C-scan SAM images were split into tiles of 300 μm × 300 μm with a shift of 20 μm in-between them. After dividing up the full-sized images, the training dataset contained 2254 different samples, where the distance between any two samples is larger than 20 μm. This is still considered as a fairly low amount of samples for DL purposes; therefore, data augmentation was used to increase the variability and number of samples. In the case of SR algorithms, various transformations can be defined on images that represent valid, possible input-output pairs. From these operations, the following were used on every image pair.

1. Random scale changes within a ratio of 0.9 and 1.1 using bi-cubic interpolation

DOI:10.15774/PPKE.ITK.2020.007

2. Rotation around the center point with ± 0.1 radian (5.7°)

3. Horizontal and vertical flipping.

After these steps, a random region of 200 μm × 200 μm was cut out from the central 240 μm × 240 μm area. These operations were simultaneously applied both on the input (180 MHz) and reference (316 MHz) images. A low-amplitude Gaussian noise (3% intensity) was added to the input image, but not to the reference image. Altogether, 32 million samples were seen by the network during the training procedure.

The parameters of the scale change, random rotation, random crop, and horizontal/vertical mirroring were generated from a uniform distribution. The transformation of the images was part of the processing pipeline and was implemented in TensorFlow version 1.12, which is an open source machine learning framework under the Apache 2.0 license [199], and was executed on two NVIDIA Tesla K80 GPUs, on which training time took 6 h. Applying these steps on every image pair independently, it could be ensured that the network has seen different samples at each step. It has to be noted that the transformed and (therefore) generated samples are not completely independent of each other, but the response of a convolutional network is invariant only for shift [200], and all these transformations result in different input images (different scale and orientation), ensuring that additional information is provided during the training of the network.

A convolutional neural network inspired by the U-Net architecture was used (see Fig. 5.1), which was originally designed for segmentation of biological samples [201], but can also be used for SR [184]. The network is a four-layered convolutional network with 16, 32, 64 and 128 features in each layer. Smooth L1 loss was selected as loss function for network training, which is a modified version of the Huber loss [202], and combines the advantages of L1 and L2 losses. For the training process, ADAm optimizer was used with the following parameters: initial learning of 10^{-4}, beta1 of 0.9 and beta2 of 0.999. For the sake of reproducibility, our implementation in TensorFlow containing all details and parameters can be found at the following link: https://github.com/horan85/acousticmicroscope. After the training procedure

DOI:10.15774/PPKE.ITK.2020.007

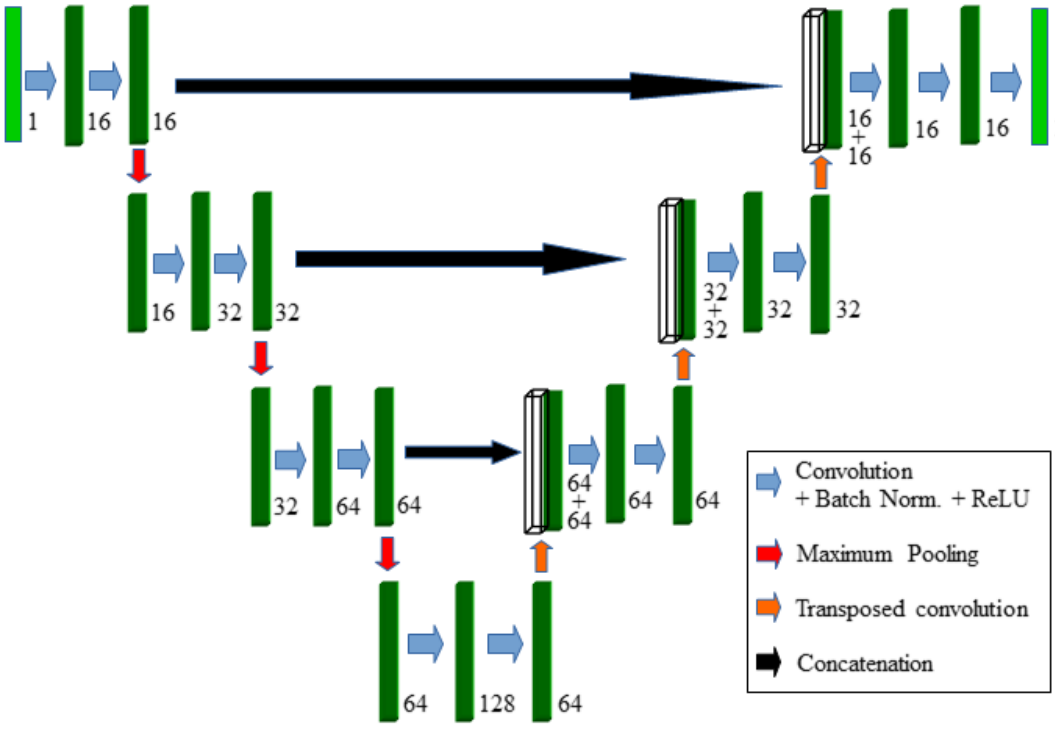

Figure 5.1: Depiction of the four-layered convolutional network, showing the order of convolution, pooling, transposed convolution, and concatenation operators. This architecture was inspired by U-Net, and as it can be seen, it processes the input image in four different scales. The left-hand side executes convolutions and downscales the results, using maximum pooling with 2 × 2 windows, while the right-hand side of the network generates the output image from the downscaled version using 3 × 3 transposed convolutions. The generated features with the same resolution are concatenated to each other. The number at each convolution block represents the number of feature maps for that convolution operator. [Th3]

was carried out, the network is ready to produce a high-resolution output image from a low-resolution input image. The output of the network is an image having a size of 200 μm × 200 μm; therefore, the test image pair was divided up into tiles of this size for quantitative evaluation, having a shift of 20 μm in-between tiles.

5.2.3 Resolution metrics

Taking the 316-MHz image as a reference, different image similarity metrics [177, 184] were used to compare the performance of resolution enhancement. The normalized root mean square error (NRMSE) is a widely used metric for quantitative

DOI:10.15774/PPKE.ITK.2020.007

comparison of two images, which calculates the MSE (see also Eq. (2.71))

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 \tag{5.1}$$

of the pixel differences between images x and y, normalized by the dynamic range of the reference image:

$$\mathrm{NRMSE} = \frac{\sqrt{\mathrm{MSE}}}{y_{max} - y_{min}}\,, \tag{5.2}$$

where y_{max} and y_{min} are the maximum and minimum values of the reference image y.

For the peak SNR (PSNR) metric, the dynamic range (or often referred to as the peak value if $y_{min} = 0$) is divided by the MSE:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{peakval^2}{\mathrm{MSE}}\right). \tag{5.3}$$

The PSNR is often used on a logarithmic scale for easier comparison, and it is also adapted in this work. The metric values were calculated as follows: based on the output of the DL method, the corresponding 200 μm × 200 μm area was selected from the original 180 MHz, Wiener, and TV images and compared to the same area of the ground truth 316-MHz image. By calculating the similarity between all the tiles and the reference image, it was possible to estimate the mean and standard deviation of PSNR and NRMSE values.

To quantify the observed resolution of the images, the following procedure was used: the FWHM of the 2-D AC function of the C-scan images (both the original and resolution enhanced) was calculated in both directions, whose average value was used as an estimation of the resolution. Note that this value is not directly comparable to the lateral resolution.

5.3 Results and discussion

The current results are presented as follows: first, the image pairs (180 MHz and 316 MHz) from the training set are shown and discussed. Next, the results on the whole test set are shown for qualitative comparison. Then, representative 200 μm × 200 μm image tiles are shown to allow further qualitative evaluation of the

DOI:10.15774/PPKE.ITK.2020.007

different image resolution enhancement techniques. For a quantitative comparison, the estimated resolution of the representative C-scan sections is included. Finally, the mean and standard deviation of the NRMSE and PSNR values over all the image tiles are presented.

We note that during the comparison of the resolution enhancement techniques, the following abbreviations are used in the figures: 180 and 316 MHz mean the original 180-MHz C-scan SAM image before any image resolution enhancement techniques applied and the reference 316-MHz C-scan SAM image, respectively. TV, Wiener, and DL stand for the methods described in Section 5.2.2.

Figure 5.2 shows the co-registered image pairs of rat and mouse brain sections taken at 180 and 316 MHz, respectively. As it can be qualitatively seen, the 316-MHz images show much better resolution with a higher level of detail. The last 180-MHz sample (4th row) became slightly contaminated during the scanning process, hence the saturated white pixels in the image.

Figure 5.3 shows the result of the different image resolution enhancement methods on the test set. Note that the DL image was constructed by stitching together the small-sized output images (see Section 5.2.2); therefore, contrast differences between patches and stitching artefacts are present. The 180- and 316-MHz images are markedly different in terms of detail, the latter having higher frequency and better resolution. The TV and the Wiener deconvolution methods show a modest improvement overall, while the result of the DL method presents the highest similarity to the reference image; however, it can also be observed that the network could not always precisely estimate the edges of the image.

To be able to further evaluate the performance of the different techniques, representative image sections taken from Fig. 5.3 (indicated by white borders) are shown in Figs. 5.4 – 5.6. In all three cases, DL is clearly seen to outperform the classical deconvolution methods, resulting in significantly improved overall sharpness, and the processed images are very similar to the 316-MHz images. However, some small scatterers visible on the 316-MHz scan disappeared during the process, which is possibly either due to the rather limited training set or the resolution limit of the initial 180-MHz scan. Further evaluation should be done to address this concern and

DOI:10.15774/PPKE.ITK.2020.007

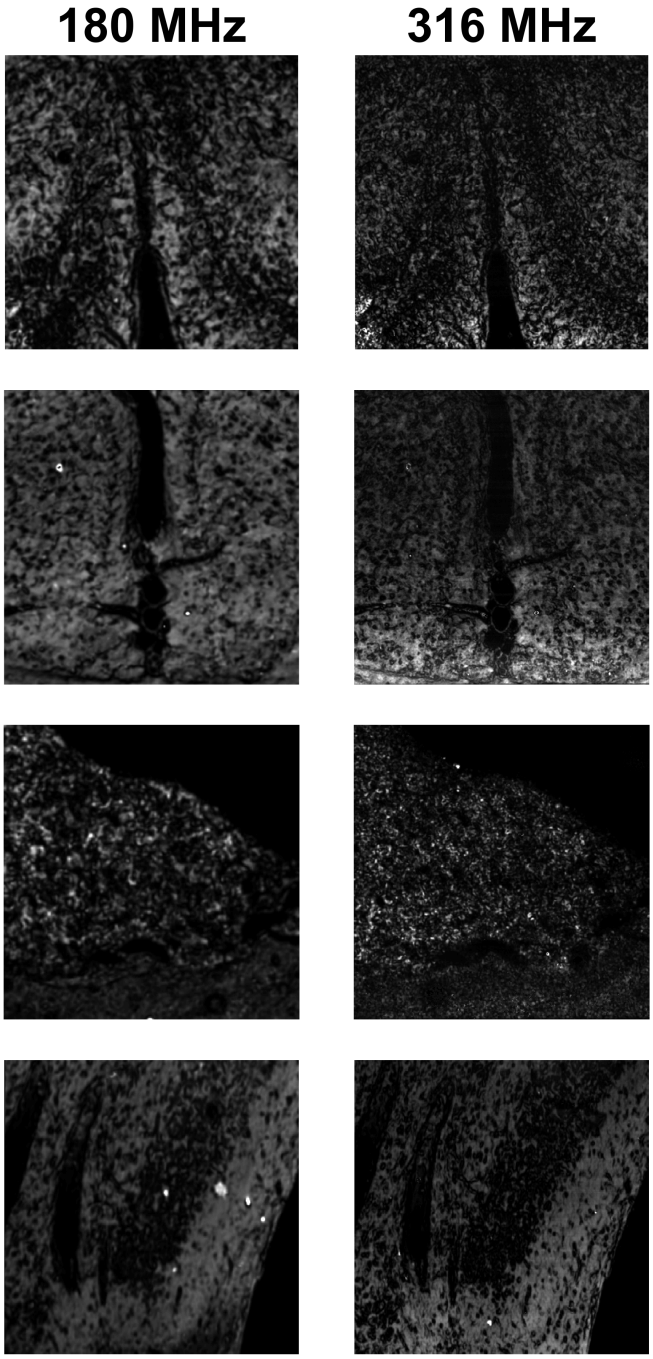

Figure 5.2: Co-registered 180- and 316-MHz C-scan SAM image pairs of rat and mouse brain sections are shown, which were split into smaller image pairs and were used as the training set. The size of the image pairs is the following (from top to bottom): 642 μm × 848 μm, 800 μm × 796 μm, 798 μm × 798 μm, 698 μm × 698 μm. [Th3]

to identify the cause. Both deconvolution techniques sometimes reveal new details compared to the original 180-MHz image. However, both methods often seem to en-

DOI:10.15774/PPKE.ITK.2020.007

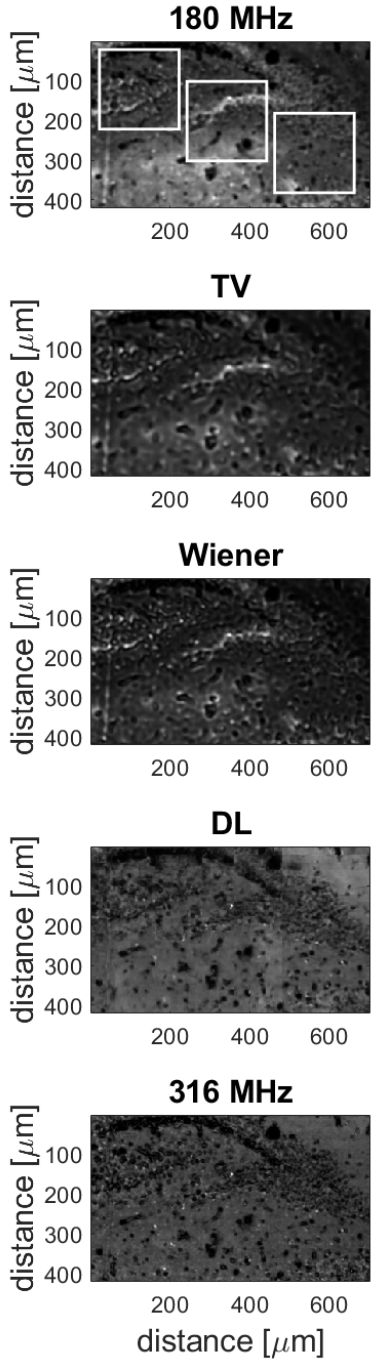

Figure 5.3: Results of the different resolution enhancement methods on the test image. The images show a rat brain coronal section (Bregma -3.12, the dentate gyrus). From top to bottom: the original 180 MHz image, slice-by-slice TV and Wiener deconvolution methods, DL and the ground truth (316 MHz) image. The areas indicated by the white borders are shown in greater detail in Figs. 5.4 – 5.6). Note the stitching artefacts present in the DL image (Section 5.2.2). [Th3]

large features, especially the TV method, where the algorithm favors homogeneous and contiguous parts. Stemming from this preference, it can also be observed that

DOI:10.15774/PPKE.ITK.2020.007

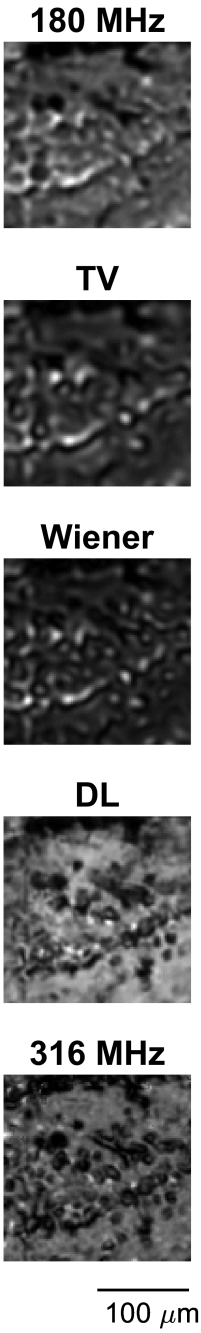

Figure 5.4: Representative sample from Fig. 5.3 (top left marked area), showing the hilus. The DL method is seen to qualitatively outperform the classical deconvolution methods in approximating the high-resolution (316 MHz) reference image. [Th3]

the Wiener method better preserves small variations in pixel values compared to the TV method. Finally, for highly variable image regions (usually densely populated with cells and vasculature), the DL method is able to provide additional details compared to the input image, returning images that are highly similar to the reference

DOI:10.15774/PPKE.ITK.2020.007

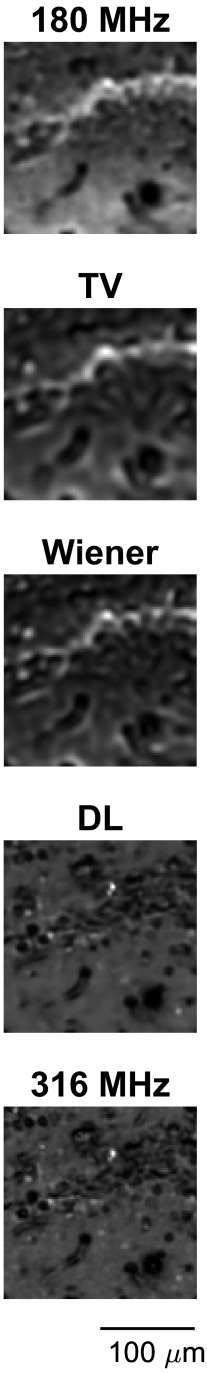

Figure 5.5: Representative sample from Fig. 5.3 (marked area in the middle), showing the lower blade of the dentate gyrus. The DL method shows a much higher qualitative similarity to the ground truth than the result of any of the classical deconvolution methods. [Th3]

image.

Table 5.2 shows the estimated resolution of the C-scan images (see Figs. 5.4 – 5.6). In general, it can be seen that the qualitative observations are verified by

DOI:10.15774/PPKE.ITK.2020.007

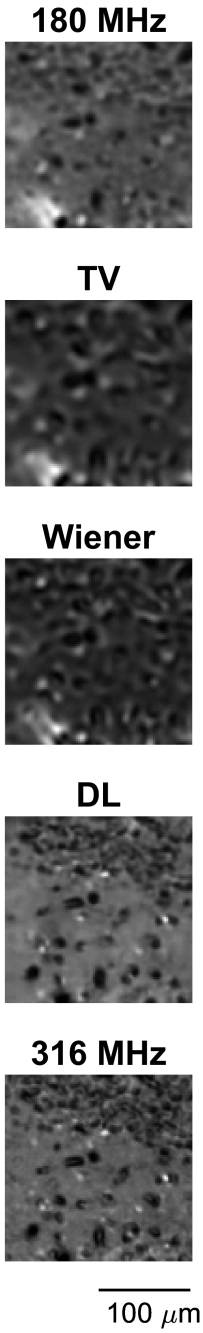

Figure 5.6: Representative sample from Fig. 5.3 (bottom right marked area), showing the neighboring thalamic nucleus. The results show that the DL method clearly outperforms both of the classical deconvolution methods. [Th3]

the quantitative values: the reference 316-MHz image having the best resolution, while the DL method being the best among the different resolution enhancement techniques, followed by the Wiener and TV methods, lastly the original unprocessed 180-MHz image. Interestingly, this approach fails to reflect the observable resolution

DOI:10.15774/PPKE.ITK.2020.007

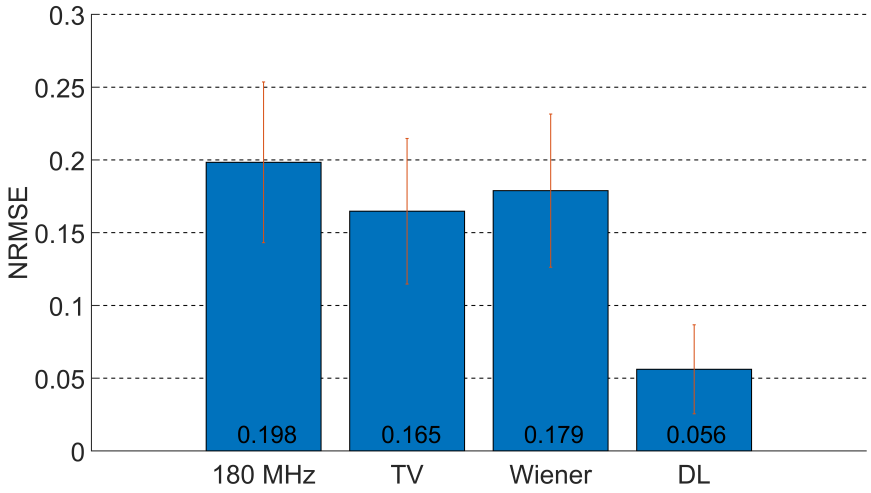

Figure 5.7: NRMSE values of the different image resolution enhancement methods (the red vertical lines showing ± 1 standard deviation). The images from the resolution enhancement methods were compared to the ground truth data (316 MHz). The values indicate an average considering all of the tiles. The DL method outperformed both the original 180-MHz image and the deconvolution methods. The TV and Wiener deconvolution methods show similar performance to each other, with a slight improvement over the original 180-MHz image. [Th3]

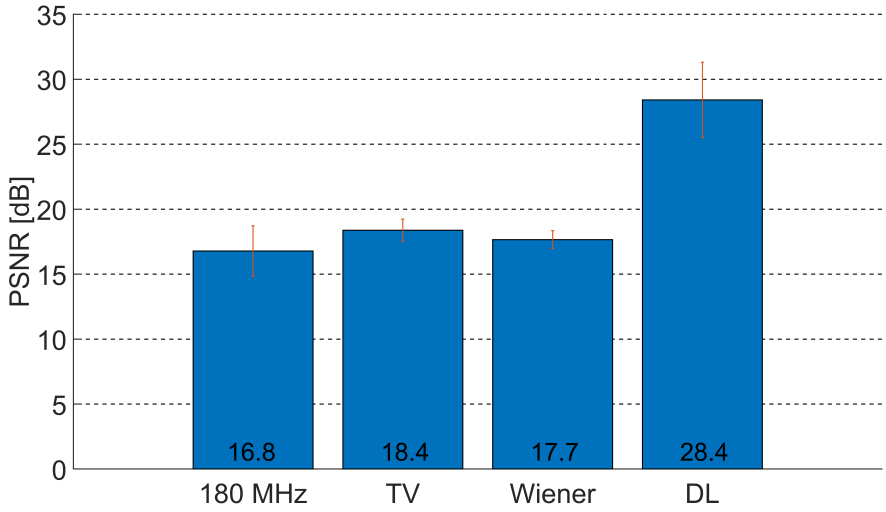

Figure 5.8: PSNR values of the different image resolution enhancement methods (the red vertical lines showing ± 1 standard deviation). The images from the resolution enhancement methods were compared to the ground truth data (316 MHz). The values indicate an average considering all of the tiles. The DL method outperformed both the original 180-MHz image and the deconvolution methods. The TV and Wiener deconvolution methods show similar performance to each other, with a slight improvement over the original 180-MHz image. [Th3]

DOI:10.15774/PPKE.ITK.2020.007

improvement in Fig. 5.4, since it suggests that all the methods (even the original 180-MHz image) outperformed the DL method.

Table 5.2: Estimated resolution of the C-scan images of Figs. 5.4 – 5.6.

Estimated resolution limit (μm)	Fig. 5.4	Fig. 5.5	Fig. 5.6
Raw 180 MHz	14.1	28.1	20.2
TV	13.5	19.7	19.3
Wiener	11.3	15.6	14.1
DL	16.0	12.4	12.4
Raw 316 MHz	11.2	11.0	10.5

Figures 5.7 and 5.8 show the average quantitative metrics over all the tiles. The results confirm the qualitative observations of Figs. 5.4 – 5.6, namely that DL greatly outperforms TV and Wiener deconvolution methods, while both show higher similarity to the reference image than the original one, with a quantitative performance being comparable to each other. The average NRMSE value of the DL method (see Fig. 5.7) was found to be 0.056, approximately a third of that of the deconvolution methods. The DL result also shows a low standard deviation in contrast to the two other techniques, demonstrating its consistent performance. The average PSNR level of the DL technique (see Fig. 5.8) is 10 dB higher than the result of the other two methods, also demonstrating the superiority of the DL method.

Although the experiments clearly demonstrate how deep learning can be used to increase the resolution (thus, decrease the resolution limit) and quality of rat and mouse brain tissue images, in every application area, especially in medical imaging, careful validation and evaluation is needed before the application of any machine learning method. In general, such approaches are evaluated using large and independent test sets and one would need such a set of other tissues to examine the performance of our approach. Even if the method is not generalizable or the method would not bring the required accuracy levels, the training set can be extended or transfer learning [203] could be used to fine-tune the method to other tissue types to increase the robustness of the procedure.

DOI:10.15774/PPKE.ITK.2020.007

Determinants of ultrasound image formation of the central nervous system (*i.e.*, surface junction of tissue components and/or cellular compartments with differing stiffness) are highly conserved among mammals. The difference, which appears from the simpler to more complex nervous system manifests as higher number and bigger size of certain cells and more numerous branching of cellular processes making possible to establish more complex neuronal networks. Thus, our assumption is that our approach could be applied and work well for tissue sections of most mammalian brains, but it is fairly difficult to estimate the level of accuracy without a proper validation set. However, drastically different samples (*e.g.*, muscle tissue or bone tissue) would require inputs in the training set describing the completely different cellular components and extracellular matrix content of these tissues, as well as their specific vascular supply with different density and 3-D arrangements. Another interesting approach would be to include optical images in the training set and see to what extent histology images could be predicted from SAM images (and vice versa).

As the results demonstrate, the intensity values vary greatly between the processed images. The aim of this work was to approximate the structure of the high-resolution image as accurately as possible and for this, the importance of intensity values were sacrificed as all activations in the neural network were normalized (also input and output samples were normalized, such as batch-normalization [204]) in the artificial neural net. In the case of quantitative value mapping, one could train a network without normalization, which would result in inferior quality regarding the structures but keep the intensity of the pixels.

To the best of our knowledge, there is no known theoretical performance limit in the DL sense if the architecture is selected properly, only input-output pairs are needed. Other properties like penetration depth, device price, and scanning time, could be limiting factors. It is also possible to produce images of multiple different resolutions using DL [205].

DOI:10.15774/PPKE.ITK.2020.007

5.4 Conclusions

This work compared two classical deconvolution-based and a DL-based image resolution enhancement method for SAM images, using a high-frequency SAM image as ground truth to evaluate the performance of the techniques.

Previous research has shown the ability of DL to perform image resolution enhancement on biomedical images [184–186]. In this work, it was shown that using even a relatively limited training set, DL greatly outperforms two common classical deconvolution techniques (Wiener and TV) and can closely approximate the reference image. At the time of publication of the corresponding article, the work seemed to be the first instance of DL being applied to improve SAM lateral resolution and resolution enhancement evaluated using experimental ground truth data. Nearly a month later, the work of Mamou *et al.* [206] had been made available online, in which they carried out similar work on SAM images and also showed how DL can be used to improve SAM resolution.

Future work could focus on training the DL neural networks on a bigger and more varied data set, as well as extending its use to 3-D data.

DOI:10.15774/PPKE.ITK.2020.007

Chapter 6

Summarizing conclusions

As a final section, the new scientific results introduced in this work are summarized in the form of thesis points.

Thesis I: *I have created an experimental method to assess the accuracy of a shift-invariant convolution-based ultrasound image formation model. The method relies on a planar arrangement of micrometer-scale scatterers in the imaging plane of a linear array. Using the coefficient of determination* R^2 *to estimate image similarity, the agreement between simulated and real images was* $R^2 = 0.43$ *for the RF image and* $R^2 = 0.65$ *for the envelope-detected B-mode image.*
Corresponding publication: [Th1]

Models of ultrasound image formation describe the forward process of how an ultrasound image is formed from an acoustic medium. Such models can be used to generate simulated ultrasound images or to obtain quantitative descriptors of the medium from real ultrasound images. A relatively simple and widely used model of image formation treats the ultrasound image (before envelope detection and compression) as the shift-invariant convolution of the imaging system point spread function (PSF) with the scattering function (SF) of the medium [40, 129].

Therefore, I created an experimental method to assess the accuracy of the convolution model. Simulated and real US images were compared to each other. The coefficient of determination was calculated both for the RF ultrasound images and the envelope-detected (B-mode) images.

DOI:10.15774/PPKE.ITK.2020.007

Various estimates of SF, PSF were tested to see which yielded the best simulation result. The source of simulation error was also explored, which possibly originates from scattering of the polystyrene particles from multiple reflections, or from microbubbles. From the observations, it is expected that by increasing the concentration of imaged scatterers or by more careful experimental design, higher overall values of the coefficient of determination can be obtained.

The results underline that, at least for the experimental setup used in the current work, the shift-invariant convolution model describes most of the variation in a B-mode image; however, care should be taken to reduce other sources of scattering such as multiple reflections or microbubbles.

Thesis II: *I have presented a novel resolution enhancement technique based on frequency-weighted axial filtering for ultrasound images that can function even when the point-spread function is shift-variant. Estimating resolution using the full-width at half maximum of the autocorrelation, the axial-lateral resolution cell was always improved, with area decreases in the range of 22–94%.*

Corresponding publication: [Th2]

Enhancement of image resolution of ultrasound images is key to help clinicians in finding early indicators of pathological lesions among others. However, the degree of improvement greatly depends on accurately estimating the PSF of the system, which in most cases is spatially variant, thus complicating its approximation and subsequent use in deconvolution.

Therefore, I investigated the possibility of using a method for US images, which is unaffected by depth-dependent effects, and it is also capable of improving the resolution both in the lateral and axial directions. Two simulated and two experimental data sets were used. The nominal central frequencies of the single-element transducers were 20 and 35 MHz. Two different deconvolution methods were used: the classical Wiener filter approach and a custom Fourier domain method (RAMP), where the signal energy was boosted with a gradually increasing function at those (higher) frequencies, where the ultrasound transducer has a weaker response. Both of the methods were used along every A-line separately. The observed resolution

DOI:10.15774/PPKE.ITK.2020.007

was quantified as the FWHM of the mean AC curves. The results confirm that frequency-weighted axial filtering can balance the need for axial and lateral resolution improvement based on their relative values with properly set parameters.

Thesis III: *I have shown the successful use of deep learning to enhance scanning acoustic microscope image lateral resolution, even with a very limited data set consisting of rat and mouse brain samples (four images in the training set, each smaller than 1 mm × 1 mm). The estimated images can closely approximate the ground truth data, having an average NRMSE of 0.056, and PSNR of 28.4 dB.*
Corresponding publication: [Th3]

Deep learning is more and more popular nowadays, yet there is limited research about its use on US images, and even those are mostly used for segmentation and classification.

Therefore, I investigated 30-μm-thick rat and mouse brain samples with a high-frequency SAM setup (180 and 316 MHz). The initial training set included 4 full size image pairs, which were co-registered. To create a properly sized training set the full-sized C-scan SAM images were split into tiles of 300 μm × 300 μm with a shift of 20 μm in-between them. Data augmentation was used to increase the variability and number of samples. A U-Net inspired neural network was used to estimate the high-resolution image based on the low-resolution image, and the 316-MHz data was used as ground truth for quantitative evaluation. Despite the training set being very limited, the results confirm the feasibility of using DL as a single-image SR method to enhance the lateral resolution of SAM images, which greatly outperformed two classical deconvolution methods (TV and Wiener deconvolution).

DOI:10.15774/PPKE.ITK.2020.007

Publications related to the thesis

[Th1] M. Gyöngy and Á. Makra, "Experimental validation of a convolution-based ultrasound image formation model using a planar arrangement of micrometer-scale scatterers," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, no. 6, pp. 1211–1219, 2015. (Cited on page(s): xiv, 40, 49, 52, 56, 57, 58, 59, 60, 61, 75, 92)

[Th2] Á. Makra, G. Csány, K. Szalai, and M. Gyöngy, "Simultaneous enhancement of B-mode axial and lateral resolution using axial deconvolution," *Proceedings of Meetings on Acoustics*, vol. 32, no. 1, 2018. (Cited on page(s): xiv, 40, 41, 43, 68, 69, 70, 71, 72, 74, 93)

[Th3] Á. Makra, W. Bost, I. Kalló, A. Horváth, M. Fournelle, and M. Gyöngy, "Enhancement of acoustic microscopy lateral resolution: A comparison between deep learning and two deconvolution methods," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 1, pp. 136–145, 2020. (Cited on page(s): xiv, 80, 83, 84, 85, 86, 87, 88, 94)

DOI:10.15774/PPKE.ITK.2020.007

Other publications of the author

[Au1] Á. Makra, "Experimental validation of an ultrasound image formation model," Bachelor's Thesis, Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, 2013. (Cited on page(s): xiv)

[Au2] Á. Makra, "An overview of sparsity-based super-resolution algorithms for medical images," in *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics 11*, G. Prószéky and P. Szolgay, Eds. Budapest, Hungary: Pázmány University ePress, 2016, pp. 161 – 164. (Cited on page(s): xiv)

[Au3] Á. Makra, "Design of a rapid scanning acoustic microscope platform for super-resolution research," in *PhD Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics 11*, G. Prószéky and P. Szolgay, Eds. Budapest, Hungary: Pázmány University ePress, 2017, pp. 49 – 49. (Cited on page(s): xiv)

[Au4] Á. Makra, "Scanning acoustic microscope system for examining biological tissue," Master's Thesis, Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, 2015. (Cited on page(s): xiv, 14, 15, 66)

[Au5] Á. Makra, J. Hatvani, and M. Gyöngy., "Calculation of equivalent ultrasound scatterers using a time-domain method," *Jedlik Laboratories Reports*, vol. 3, no. JLR/3-2015, pp. 7 – 12, 2015. (Cited on page(s): xiv)

[Au6] K. Füzesi, Á. Makra, and M. Gyöngy, "A stippling algorithm to generate equivalent point scatterer distributions from ultrasound images," in *Proceed-*

DOI:10.15774/PPKE.ITK.2020.007

ings of Meetings on Acoustics 6ICU, vol. 32, no. 1. ASA, 2017, p. 020008. (Cited on page(s): 40)

DOI:10.15774/PPKE.ITK.2020.007

Bibliography

[1] M. Gyöngy, "Estimation of PSF from an ultrasound image," Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Tech. Rep., 2017. (Cited on page(s): xiv)

[2] N. S. Berko, J. N. Le, B. A. Thornhill, D. Wang, A. Negassa, E. S. Amis, and M. Koenigsberg, "Design and validation of a peer-teacher-based musculoskeletal ultrasound curriculum," *Academic Radiology*, vol. 26, no. 5, pp. 701–706, 2019. (Cited on page(s): 1)

[3] J. A. Hides, D. H. Cooper, and M. J. Stokes, "Diagnostic ultrasound imaging for measurement of the lumbar multifidus muscle in normal young adults," *Physiotherapy Theory and Practice*, vol. 8, no. 1, pp. 19–26, 1992. (Cited on page(s): 1)

[4] R. Coelho, H. Ribeiro, and G. Maconi, "Bowel thickening in crohn's disease," *Inflammatory Bowel Diseases*, vol. 23, no. 1, pp. 23–34, 2017. (Cited on page(s): 1)

[5] G. T. Haar, "Ultrasound focal beam surgery," *Ultrasound in Medicine & Biology*, vol. 21, no. 9, pp. 1089 – 1100, 1995. (Cited on page(s): 1)

[6] F. Sammartino, D. W. Beam, J. Snell, and V. Krishna, "Kranion, an open-source environment for planning transcranial focused ultrasound surgery: technical note," *Journal of Neurosurgery*, pp. 1–7, 2019. (Cited on page(s): 1)

[7] W. She, T. Cheung, C. R. Jenkins, and M. G. Irwin, "Clinical applications of high-intensity focused ultrasound," *Hong Kong Medical Journal*, vol. 22, no. 4, pp. 382 – 392, 2016. (Cited on page(s): 1)

DOI:10.15774/PPKE.ITK.2020.007

[8] A. Iula, "Ultrasound systems for biometric recognition," *Sensors*, vol. 19, no. 10, p. 2317, 2019. (Cited on page(s): 1)

[9] M. Kersemans, E. Verboven, J. Segers, S. Hedayatrasa, and W. V. Paepegem, "Non-destructive testing of composites by ultrasound, local defect resonance and thermography," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 2, no. 8, 2018, p. 554. (Cited on page(s): 1, 74)

[10] Z. Remili, Y. Ousten, B. Levrier, E. Suhir, and L. Bechou, "Scanning acoustic microscopy and shear wave imaging mode performances," *IEEE 65th Electronic Components and Technology Conference (ECTC)*, pp. 2090–2101, 26-29 May 2015. (Cited on page(s): 1, 74)

[11] A. Phommahaxay, I. D. Wolf, T. Duric, P. Hoffrogge, S. Brand, P. Czurratis, H. Philipsen, G. Beyer, H. Struyf, and E. Beyne, "Defect detection in through silicon vias by GHz scanning acoustic microscopy: key ultrasonic characteristics," *IEEE 64th Electronic Components and Technology Conference (ECTC),*, pp. 850–855, 27-30 May 2014. (Cited on page(s): 1, 74)

[12] M. Fan, L. Su, L. Li, W. Wei, Z. He, C. Wong, and X. Lu, "A fuzzy SVM for intelligent diagnosis of solder bumps using scanning acoustic microscopy," *Semiconductor Technology International Conference (CSTIC)*, 13-14 March 2016. (Cited on page(s): 1, 74)

[13] F. Naumann and S. Brand, "Numerical prototyping and defect evaluation of scanning acoustic microscopy for advanced failure diagnostics," *17th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems*, pp. 1–7, 18-20 April 2016. (Cited on page(s): 1, 74)

[14] S. Brand, F. Naumann, S. Tismer, B. Boettge, J. Rudzki, F. Osterwald, and M. Petzold, "Non-destructive assessment of reliability and quality related properties of power electronic devices for the in-line application of scanning acoustic microscopy," *9th International Conference on Integrated Power Electronics Systems (CIPS)*, 8-10 March 2016. (Cited on page(s): 1, 74)

DOI:10.15774/PPKE.ITK.2020.007

[15] S. Brand, E. C. Weiss, R. M. Lemor, and M. C. Kolios, "High frequency ultrasound tissue characterization and acoustic microscopy of intracellular changes," *Ultrasound in Medicine and Biology*, vol. 34, pp. 1396–1407, 2008. (Cited on page(s): 1, 74)

[16] E. Grünwald, R. Hammer, R. Jördis, B. Sartory, and R. Brunner, "Accretion detection via scanning acoustic microscopy in microelectronic components-considering symmetry breaking effects," *Microscopy and Microanalysis 23*, vol. 23, pp. 1466–1467, 2017. (Cited on page(s): 1, 74)

[17] M. Kim, N. Choi, Y. I. Kim, and Y. H. Lee, "Characterization of RF sputtered zinc oxide thin films on silicon using scanning acoustic microscopy," *Journal of Electroceramics*, pp. 1–9, 2017. (Cited on page(s): 1, 74)

[18] D. Wang, X. He, Z. Xu, W. Jiao, F. Yang, L. Jiang, L. Li, W. Liu, and R. Wang, "Study on damage evaluation and machinability of UD-CFRP for the orthogonal cutting operation using scanning acoustic microscopy and the finite element method," *Materials*, vol. 10, p. 204, 2017. (Cited on page(s): 1, 74)

[19] J. Dong, X. Wu, A. Locquet, and D. S. Citrin, "Terahertz superresolution stratigraphic characterization of multilayered structures using sparse deconvolution," *IEEE TRANSACTIONS ON TERAHERTZ SCIENCE AND TECHNOLOGY*, vol. 7, pp. 260–267, 2017. (Cited on page(s): 1, 74)

[20] M. S. Firouz, A. Farahmandi, and S. Hosseinpour, "Recent advances in ultrasound application as a novel technique in analysis, processing and quality control of fruits, juices and dairy products industries: A review," *Ultrasonics Sonochemistry*, vol. 57, pp. 73–88, 2019. (Cited on page(s): 1)

[21] N. Segura, M. Amarillo, N. Martinez, M. Grompone *et al.*, "Improvement in the extraction of hass avocado virgin oil by ultrasound application," *J. Food Res*, vol. 7, pp. 106–113, 2018. (Cited on page(s): 1)

DOI:10.15774/PPKE.ITK.2020.007

[22] J. M. del Fresno, I. Loira, A. Morata, C. González, J. A. Suárez-Lepe, and R. Cuerda, "Application of ultrasound to improve lees ageing processes in red wines," *Food Chemistry*, vol. 261, pp. 157–163, 2018. (Cited on page(s): 1)

[23] V. Akdeniz and A. S. Akalın, "New approach for yoghurt and ice cream production: High-intensity ultrasound," *Trends in Food Science & Technology*, vol. 86, pp. 392 – 398, 2019. (Cited on page(s): 1)

[24] D. Rohrbach, B. Wodlinger, J. Wen, J. Mamou, and E. Feleppa, "High-frequency quantitative ultrasound for imaging prostate cancer using a novel micro-ultrasound scanner," *Ultrasound in Medicine & Biology*, vol. 44, no. 7, pp. 1341–1354, 2018. (Cited on page(s): 1)

[25] T. Mizoguchi, K. Tamura, J. Mamou, J. A. Ketterling, K. Yoshida, and T. Yamaguchi, "Comprehensive backscattering characteristics analysis for quantitative ultrasound with an annular array: a basic study on homogeneous scattering phantom," *Japanese Journal of Applied Physics*, vol. 58, no. SG, p. SGGE08, 2019. (Cited on page(s): 1)

[26] M. L. Oelze and J. Mamou, "Review of quantitative ultrasound: Envelope statistics and backscatter coefficient imaging and contributions to diagnostic ultrasound," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 63, no. 2, pp. 336 – 351, 2016. (Cited on page(s): 1)

[27] E. J. Feleppa, J. Mamou, and D. Rohrbach, "Typing and imaging of biological and non-biological materials using quantitative ultrasound," 2019, uS Patent 10,338,033. (Cited on page(s): 1)

[28] J. Mamou, P. Goundan, D. Rohrbach, H. Patel, E. Feleppa, and S. Lee, "In-vivo-quantitative-ultrasound assessment of thyroid nodules," *The Journal of the Acoustical Society of America*, vol. 146, no. 4, pp. 2811 – 2812, 2019. (Cited on page(s): 1)

[29] S. C. Lin, E. Heba, T. Wolfson, B. Ang, A. Gamst, A. Han, J. W. Erdman, W. D. O'Brien, M. P. Andre, C. B. Sirlin, and R. Loomba, "Noninvasive diagnosis

DOI:10.15774/PPKE.ITK.2020.007

of nonalcoholic fatty liver disease and quantification of liver fat using a new quantitative ultrasound technique," *Clinical Gastroenterology and Hepatology*, vol. 13, no. 7, pp. 1337–1345.e6, 2015. (Cited on page(s): 1)

[30] L. C. Slane, J. Martin, R. DeWall, D. Thelen, and K. Lee, "Quantitative ultrasound mapping of regional variations in shear wave speeds of the aging achilles tendon," *European Radiology*, vol. 27, no. 2, pp. 474–482, 2016. (Cited on page(s): 1)

[31] N. G. Simon, J. W. Ralph, C. Lomen-Hoerth, A. N. Poncelet, S. Vucic, M. C. Kiernan, and M. Kliot, "Quantitative ultrasound of denervated hand muscles," *Muscle & Nerve*, vol. 52, no. 2, pp. 221–230, 2015. (Cited on page(s): 1)

[32] J. S. Paige, G. S. Bernstein, E. Heba, E. A. Costa, M. Fereirra, T. Wolfson, A. C. Gamst, M. A. Valasek, G. Y. Lin, A. Han *et al.*, "A pilot comparative study of quantitative ultrasound, conventional ultrasound, and mri for predicting histology-determined steatosis grade in adult nonalcoholic fatty liver disease," *American Journal of Roentgenology*, vol. 208, no. 5, pp. W168–W177, 2017. (Cited on page(s): 1)

[33] J. Mamou, D. Rohrbach, E. Saegusa-Beecroft, E. Yanagihara, J. Machi, and E. J. Feleppa, "Ultrasound-scattering models based on quantitative acoustic microscopy of fresh samples and unstained fixed sections from cancerous human lymph nodes," in *2015 IEEE International Ultrasonics Symposium (IUS)*. IEEE, 2015. (Cited on page(s): 1)

[34] J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue," *J. Acoust. Soc. Am.*, vol. 89, no. 1, pp. 182–190, 1991. (Cited on page(s): 2, 17, 25, 28, 38, 40, 46, 47, 48)

[35] R. Zemp, C. Abbey, and M. Insana, "Linear system models for ultrasonic imaging: application to signal statistics," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 50, no. 6, pp. 642–654, 2003. (Cited on page(s): 2, 28, 38, 46)

DOI:10.15774/PPKE.ITK.2020.007

[36] J. P. Kemmerer, M. L. Oelze, and M. Gyöngy, "Scattering by single physically large and weak scatterers in the beam of a single-element transducer," *The Journal of the Acoustical Society of America*, vol. 137, no. 3, pp. 1153–1163, 2015. (Cited on page(s): 2, 47, 51)

[37] R. S. C. Cobbold, *Foundations of Biomedical Ultrasound*. Oxford University Press, 2007. (Cited on page(s): 2, 6, 10, 11, 17, 23, 25, 26, 27, 41)

[38] J. M. Mari, T. Blu, O. B. Matar, M. Unser, and C. Cachard, "A bulk modulus dependent linear model for acoustical imaging," *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2413–2419, 2009. (Cited on page(s): 2, 25, 38, 40, 46)

[39] J. Mamou, M. L. Oelze, W. D. O'Brien Jr., and J. F. Zachary, "Extended three-dimensional impedance map methods for identifying ultrasonic scattering sites," *J. Acoust. Soc. Am.*, vol. 123, no. 2, pp. 1195–1208, 2008. (Cited on page(s): 2, 45)

[40] J. C. Bamber and R. Dickinson, "Ultrasonic B-scanning: A computer simulation," *Phys. Med. Biol.*, vol. 25, no. 3, pp. 463–479, 1980. (Cited on page(s): 2, 44, 92)

[41] T. Hergum, S. Langeland, E. Remme, and H. Torp, "Fast ultrasound imaging simulation in K-space," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, no. 6, pp. 1159–1167, 2009. (Cited on page(s): 3, 45)

[42] P. Howell, "Resolution-enhancement method for digital imaging," 2003, uS Patent 6,570,613. (Cited on page(s): 3)

[43] A. Kumar, S. G. Eraniose, and A. K. Lakshmikumar, "Multi-resolution ip camera," 2013, uS Patent App. 13/447,202. (Cited on page(s): 3)

[44] V. A. Trofimov and V. V. Trofimov, "IR camera temperature resolution enhancing using computer processing of IR image," in *Thermosense: Thermal Infrared Applications XXXVIII*, J. N. Zalameda and P. Bison, Eds. SPIE, 2016. (Cited on page(s): 3)

DOI:10.15774/PPKE.ITK.2020.007

[45] H. Demirel and G. Anbarjafari, "Satellite image resolution enhancement using complex wavelet transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp. 123–126, 2010. (Cited on page(s): 3)

[46] O. Harikrishna and A. Maheshwari, "Satellite image resolution enhancement using dwt technique," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, no. 5, pp. 274 – 278, 2012. (Cited on page(s): 3)

[47] H. Demirel and G. Anbarjafari, "Discrete wavelet transform-based satellite image resolution enhancement," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 1997–2004, 2011. (Cited on page(s): 3)

[48] M. Z. Iqbal, A. Ghafoor, and A. M. Siddiqui, "Satellite image resolution enhancement using dual-tree complex wavelet transform and nonlocal means," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 451–455, 2013. (Cited on page(s): 3)

[49] P. Rasti, I. Lusi, H. Demirel, R. Kiefer, and G. Anbarjafari, "Wavelet transform based new interpolation technique for satellite image resolution enhancement," in *2014 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology*. IEEE. (Cited on page(s): 3)

[50] M. Farrar and E. Smith, "Spatial resolution enhancement of terrestrial features using deconvolved SSM/i microwave brightness temperatures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 2, pp. 349–355, 1992. (Cited on page(s): 3)

[51] G. Daniell and S. Gull, "Maximum entropy algorithm applied to image enhancement," *IEE Proceedings E-Computers and Digital Techniques*, vol. 127, no. 5, pp. 170–172, 1980. (Cited on page(s): 3)

[52] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," in *ACM transactions on graphics (TOG)*, vol. 23, no. 3. ACM, 2004, pp. 673–678. (Cited on page(s): 3)

DOI:10.15774/PPKE.ITK.2020.007

[53] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Computer Graphics Forum*, vol. 28, no. 1, pp. 161–171, 2009. (Cited on page(s): 3)

[54] T. Schulz, B. Stribling, and J. Miller, "Multiframe blind deconvolution with real data: imagery of the hubble space telescope," *Optics Express*, vol. 1, no. 11, p. 355, 1997. (Cited on page(s): 3)

[55] A. S. Carasso, "APEX blind deconvolution of color hubble space telescope imagery and other astronomical data," *Optical Engineering*, vol. 45, no. 10, p. 107004, 2006. (Cited on page(s): 3)

[56] J. L. Hershey, "A deconvolution technique for hubble space telescope fgs fringe analysis," *Publications of the Astronomical Society of the Pacific*, vol. 104, no. 677, pp. 592 – 596, 1992. (Cited on page(s): 3)

[57] R. L. White, "Image restoration using the damped lucy–richardson method," in *Instrumentation in Astronomy VIII*, D. L. Crawford and E. R. Craine, Eds. SPIE, 1994. (Cited on page(s): 3, 33)

[58] T. L. Szabo, *Diagnostic Ultrasound Imaging: Inside Out.* Academic Press, 2004. (Cited on page(s): 3, 6, 74)

[59] (2010) Olympus ultrasonic transducers technical notes. Last visited: 2019-12-11. [Online]. Available: http://www.olympus-ims.com/data/File/panametrics/panametrics-UT.en.pdf (Cited on page(s): 6, 8, 9, 11)

[60] J. Curie and P. Curie, "Développement par compression de l'électricité polaire dans les cristaux hémièdres à faces inclinées," *Bulletin de minéralogie*, vol. 3, no. 4, pp. 90–93, 1880. (Cited on page(s): 6)

[61] A. Manbachi and R. S. C. Cobbold, "Development and application of piezoelectric materials for ultrasound generation and detection," *Ultrasound*, vol. 19, no. 4, pp. 187–196, 2011. (Cited on page(s): 6)

[62] H. Kawai, "The piezoelectricity of poly (vinylidene fluoride)," *Japanese Journal of Applied Physics*, vol. 8, no. 7, pp. 975–976, 1969. (Cited on page(s): 6)

DOI:10.15774/PPKE.ITK.2020.007

[63] E. O. Belcher, D. C. Lynn, H. Q. Dinh, and T. J. Laughlin, "Beamforming and imaging with acoustic lenses in small, high-frequency sonars," in *Oceans' 99. MTS/IEEE. Riding the Crest into the 21st Century. Conference and Exhibition. Conference Proceedings (IEEE Cat. No. 99CH37008)*, vol. 3. IEEE, 1999, pp. 1495–1499. (Cited on page(s): 6)

[64] J. M. Cannata, T. A. Ritter, W.-H. Chen, and K. K. Shung, "Design of Focused Single Element (50 - 100 MHz) Transducers using Lithium Niobate," *IEEE Ultrasonisc Symposium*, vol. 25, no. 7, pp. 1129–1133, 1969. (Cited on page(s): 6)

[65] B. Yin, D. Xing, Y. Wang, Y. Zeng, Y. Tan, and Q. Chen, "Fast photoacoustic imaging system based on 320-element linear transducer array," *Physics in Medicine and Biology*, vol. 49, no. 7, pp. 1339–1346, 2004. (Cited on page(s): 7)

[66] Fwhm definition. Last visited: 2019-12-11. [Online]. Available: http://commons.wikimedia.org/wiki/File:FWHM.svg (Cited on page(s): 10)

[67] G. S. Kino, *Acoustic Waves: Devices, Imaging, and Analog Signal Processing*. Englewood Cliffs, NJ:Prentice Hall, 1987. (Cited on page(s): 11)

[68] P. N. Burns, "Introduction to the physical principles of ultrasound imaging and Doppler," 2005. (Cited on page(s): 12, 13)

[69] T. Chernyakova, D. Cohen, M. Shoham, and Y. C. Eldar, "iMAP beamforming for high-quality high frame rate imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 66, no. 12, pp. 1830–1844, 2019. (Cited on page(s): 15)

[70] T. L. Szabo, "Time domain wave equations for lossy media obeying a frequency power law," *The Journal of the Acoustical Society of America*, vol. 96, no. 1, pp. 491–500, jul 1994. (Cited on page(s): 16)

[71] M. Gyöngy, "Passive Cavitation Mapping for Monitoring Ultrasound Therapy," Ph.D. dissertation, University of Oxford, 2010. (Cited on page(s): 17, 22, 25)

DOI:10.15774/PPKE.ITK.2020.007

[72] F. Gran, "Spatio-temporal encoding in medical ultrasound imaging," Ph.D. dissertation, Technical University of Denmark, 2005. (Cited on page(s): 17, 18, 21, 25)

[73] N. A. Gumerov and R. Duraiswami, *Fast Multipole Methods for the Helmholtz Equation in Three Dimensions*, 1st ed. Elsevier, 2005. (Cited on page(s): 17)

[74] L. Tamás, *Az Áramlástan alapjai*. Budapest University of Technology and Economics, Department of Fluid Mechanics, 1992, előadás jegyzet. (Cited on page(s): 17, 20)

[75] P. R. Stepanishen, "Pulsed transmit/receive response of ultrasonic piezoelectric transducers," *The Journal of the Acoustical Society of America*, vol. 69, no. 6, pp. 1815–1827, 1981. (Cited on page(s): 28)

[76] T. E. Bishop, S. D. Babacan, B. Amizic, A. K. Katsaggelos, T. Chan, and R. Molina, *Blind image deconvolution: problem formulation and existing approaches*. CRC press, 2007. (Cited on page(s): 30)

[77] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2009. (Cited on page(s): 30)

[78] A. Goldstein and R. Fattal, "Blur-kernel estimation from spectral irregularities," in *European Conference on Computer Vision*. Springer, 2012, pp. 622 – 635. (Cited on page(s): 30)

[79] W. Hu, J. Xue, and N. Zheng, "PSF estimation via gradient domain correlation," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 386 – 392, 2012. (Cited on page(s): 30)

[80] R. Wang and D. Tao, "Recent progress in image deblurring," *arXiv preprint arXiv:1409.6838*, 2014. (Cited on page(s): 30)

[81] S. Chaudhuri, R. Velmurugan, and R. Rameshan, *Blind deconvolution methods: A review*. Springer, 2014. (Cited on page(s): 30)

DOI:10.15774/PPKE.ITK.2020.007

[82] R. Morin, A. Basarab, and D. Kouame, "Alternating direction method of multipliers framework for super-resolution in ultrasound imaging," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, may 2012. (Cited on page(s): 30)

[83] J. Jensen and S. Leeman, "Nonparametric estimation of ultrasound pulses," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 10, pp. 929–936, 1994. (Cited on page(s): 31, 40)

[84] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 3rd ed. Pearson Education, 2010. (Cited on page(s): 31)

[85] D. Adam and O. Michailovich, "Blind deconvolution of ultrasound sequences using nonparametric local polynomial estimates of the pulse," *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, vol. 49, pp. 118–131, 2002. (Cited on page(s): 31, 75)

[86] T. Taxt, "Comparison of cepstrum-based methods for radial blind deconvolution of ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 44, pp. 666–674, 1997. (Cited on page(s): 31, 75)

[87] J. A. Jensen, "Deconvolution of ultrasound images," *Ultrasonic imaging*, vol. 14, pp. 1 – 15, 1992. (Cited on page(s): 31, 75)

[88] R. Jirik and T. Taxt, "Two-dimensional blind bayesian deconvolution of medical ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 55, no. 10, pp. 2140 – 2153, 2008. (Cited on page(s): 31, 34)

[89] U. Abeyratne, A. Petropulu, J. Reid, T. Golas, E. Conant, and F. Forsberg, "Higher order versus second order statistics in ultrasound image deconvolution," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 44, no. 6, pp. 1409 – 1416, 1997. (Cited on page(s): 31)

[90] S. Wan, B. Raju, and M. Srinivasan, "Robust deconvolution of high-frequency ultrasound images using higher-order spectral analysis and wavelets," *IEEE*

DOI:10.15774/PPKE.ITK.2020.007

Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 50, no. 10, pp. 1286 – 1295, 2003. (Cited on page(s): 31)

[91] T. Taxt, "Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 42, pp. 543–554, 1995. (Cited on page(s): 31, 75)

[92] O. Michailovich and D. Adam, "A novel approach to the 2-d blind deconvolution problem in medical ultrasound," *IEEE Transactions on Medical Imaging*, vol. 24, no. 1, pp. 86 – 104, 2005. (Cited on page(s): 31)

[93] P. A. Jansson, *Deconvolution of images and spectra*. Courier Corporation, 2014. (Cited on page(s): 32)

[94] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. MIT Press, 1950. (Cited on page(s): 32)

[95] D. G. Lampard, "Generalization of the wiener-khintchine theorem to nonstationary processes," *Journal of Applied Physics*, vol. 25, no. 6, pp. 802–803, 1954. (Cited on page(s): 32)

[96] A. V. Oppenheim and G. C. Verghese, *Signals, Systems and Inference: Class Notes for 6.011, Introduction to Communication, Control and Signal Processing*. Massachusetts Institute of Technology: MIT OpenCourseWare, 2010, https://ocw.mit.edu. License: Creative Commons BY-NC-SA. (Cited on page(s): 32)

[97] W. H. Richardson, "Bayesian-based iterative method of image restoration," *Journal of the Optical Society of America*, vol. 62, no. 1, pp. 55 – 59, 1972. (Cited on page(s): 33)

[98] L. B. Lucy, "An iterative technique for the rectification of observed distributions," *The Astronomical Journal*, vol. 79, pp. 745 – 754, 1974. (Cited on page(s): 33)

DOI:10.15774/PPKE.ITK.2020.007

[99] D. S. C. Biggs and M. Andrews, "Acceleration of iterative image restoration algorithms," *Applied Optics*, vol. 36, no. 8, pp. 1766 – 1775, 1997. (Cited on page(s): 33)

[100] C. Dalitz, R. Pohle-Frohlich, and T. Michalk, "Point spread functions and deconvolution of ultrasonic images," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 62, no. 3, pp. 531–544, 2015. (Cited on page(s): 33)

[101] E. Boukouvala and A. H. Lettington, "Restoration of astronomical images by an iterative superresolution algorithm," *Astronomy & Astrophysics*, vol. 399, no. 3, pp. 807–811, 2003. (Cited on page(s): 33)

[102] J. Rameau, H.-B. Yang, and P. Johnson, "Application of the lucy–richardson deconvolution procedure to high resolution photoemission spectra," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 181, no. 1, pp. 35–43, 2010. (Cited on page(s): 33)

[103] E. Pantin and J.-L. Starck, "Deconvolution of astronomical images using the multiscale maximum entropy method," *Astronomy and Astrophysics Supplement Series*, vol. 118, no. 3, pp. 575–585, 1996. (Cited on page(s): 33)

[104] E. S. Meinel, "Origins of linear and nonlinear recursive restoration algorithms," *Journal of the Optical Society of America A*, vol. 3, no. 6, pp. 787 – 799, 1986. (Cited on page(s): 34)

[105] H. M. Adorf, R. N. Hook, L. B. Lucy, and F. D. Murtagh, "Accelerating the richardson–lucy restoration algorithm," in *European Southern Observatory Conference and Workshop Proceedings*, vol. 41, 1992, pp. 99–103. (Cited on page(s): 34)

[106] T. J. Holmes and Y.-H. Liu, "Acceleration of maximum-likelihood image restoration for fluorescence microscopy and other noncoherent imagery," *Journal of the Optical Society of America A*, vol. 8, no. 6, pp. 893 – 907, 1991. (Cited on page(s): 34)

DOI:10.15774/PPKE.ITK.2020.007

[107] L. Kaufman, "Implementing and accelerating the EM algorithm for positron emission tomography," *IEEE Transactions on Medical Imaging*, vol. 6, no. 1, pp. 37 – 51, 1987. (Cited on page(s): 34)

[108] D. S. C. Biggs and M. Andrews, "Conjugate gradient acceleration of maximum-likelihood image restoration," *Electronics Letters*, vol. 31, no. 23, pp. 1985–1986, 1995. (Cited on page(s): 34)

[109] R. G. Lane, "Methods for maximum-likelihood deconvolution," *Journal of the Optical Society of America A*, vol. 13, no. 10, pp. 1992 – 1998, 1996. (Cited on page(s): 34)

[110] N. Zhao, Q. Wei, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Blind deconvolution of medical ultrasound images using a parametric model for the point spread function," in *2016 IEEE International Ultrasonics Symposium (IUS)*. IEEE, 2016. (Cited on page(s): 34)

[111] N. Zhao, A. Basarab, D. Kouamé, and J.-Y. Tourneret, "Joint segmentation and deconvolution of ultrasound images using a hierarchical bayesian model based on generalized gaussian priors," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3736 – 3750, 2016. (Cited on page(s): 34)

[112] R. Morin, S. Bidon, A. Basarab, and D. Kouame, "Semi-blind deconvolution for resolution enhancement in ultrasound imaging," in *2013 IEEE International Conference on Image Processing*. IEEE, 2013. (Cited on page(s): 34)

[113] Z. Huang, J. Zhang, and C. Zhang, "Ultrasound image reconstruction by two-dimensional blind total variation deconvolution," in *2009 IEEE International Conference on Control and Automation*. IEEE, 2009. (Cited on page(s): 34)

[114] C. Yu, C. Zhang, and L. Xie, "An envelope signal based deconvolution algorithm for ultrasound imaging," *Signal Processing*, vol. 92, no. 3, pp. 793 – 800, 2012. (Cited on page(s): 34, 75)

DOI:10.15774/PPKE.ITK.2020.007

[115] C. Yu, C. Zhang, and L. Xie, "A blind deconvolution approach to ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 59, no. 2, pp. 271 – 280, 2012. (Cited on page(s): 34)

[116] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. W. L. Aerts, "Artificial intelligence in radiology," *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, 2018. (Cited on page(s): 35)

[117] K. W. Johnson, J. T. Soto, B. S. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley, and J. T. Dudley, "Artificial intelligence in cardiology," *Journal of the American College of Cardiology*, vol. 71, no. 23, pp. 2668–2679, 2018. (Cited on page(s): 35)

[118] F. Sebastiani, "Machine learning in automated text categorization," *ACM Computing Surveys*, vol. 34, no. 1, pp. 1–47, 2002. (Cited on page(s): 35)

[119] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016. (Cited on page(s): 35)

[120] K.-C. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini, and F. Menczer, "Arming the public with artificial intelligence to counter social bots," *Human Behavior and Emerging Technologies*, vol. 1, no. 1, pp. 48–61, 2019. (Cited on page(s): 35)

[121] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097–1105, 2012. (Cited on page(s): 35)

[122] A. Karpathy. Stanford CS class CS231n convolutional neural networks for visual recognition. Last visited: 2019-12-13. [Online]. Available: http://cs231n.github.io/neural-networks-1/ (Cited on page(s): 35)

[123] C. Benedek. PPKE ITK Basic image processing course lecture slides 2019/20. Last visited: 2019-12-13. [Online]. Available: https://users.itk.ppke.hu/kep/Lectures/IPA_13_Deep_Learning.pdf (Cited on page(s): 36)

[124] J. Hatvani, "SISR techniques for 3D medical images," Institut de Recherche en Informatique de Toulouse (IRIT), Tech. Rep., MINDS PhD séminair, 2019-12-10. (Cited on page(s): 37)

[125] R. G. Dantas, E. T. Costa, and S. Leeman, "Ultrasound speckle and equivalent scatterers," *Ultrasonics*, vol. 43, no. 6, pp. 405 – 420, may 2005. (Cited on page(s): 38, 39)

[126] L. Mo and R. Cobbold, "Theoretical models of ultrasonic scattering in blood," *Ultrasonic scattering in biological tissues*, pp. 125 – 170, 1993. (Cited on page(s): 40)

[127] K. Füzesi, "Tissue characterization and custom manufactured phantoms for modelling of medical ultrasound images," Ph.D. dissertation, Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, Roska Tamás Doctoral School of Sciences and Technology, 2019. (Cited on page(s): 40)

[128] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, no. 2, pp. 262–267, 1992. (Cited on page(s): 40, 47, 53, 77)

[129] C. R. Hill, J. C. Bamber, and G. R. ter Haar, *Physical Principles of Medical Ultrasonics*, 2nd ed. Chichester, UK:John Wiley and Sons, 2004. (Cited on page(s): 44, 92)

[130] R. M. Cramblitt and K. J. Parker, "Generation of non-Rayleigh speckle distributions using marked regularity models," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 46, no. 4, pp. 867–874, 1999. (Cited on page(s): 45)

DOI:10.15774/PPKE.ITK.2020.007

[131] E. M. G. P. Jacobs and J. M. Thijssen, "A simulation study of echographic imaging of diffuse and structurally scattering media," *Ultrason. Imaging*, vol. 13, no. 4, pp. 316–333, 1991. (Cited on page(s): 45)

[132] R. C. Waag, J. O. Nilsson, and J. P. Astheimer, "Characterization of volume scattering power spectra in isotropic media from power spectra of scattering by planes," *J. Acoust. Soc. Am.*, vol. 74, no. 5, pp. 1555–1571, 1983. (Cited on page(s): 45, 47)

[133] F. L. Lizzi, M. C. Rorke, D. L. King, E. J. Feleppa, A. Kalisz, J. B. Sokil-Melgar, and J. Nowakowski, "Simulation studies of ultrasonic backscattering and B-mode images of liver using acoustic microscopy data," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, no. 2, pp. 212–226, 1992. (Cited on page(s): 45)

[134] A. J. Dapore, M. R. King, J. Harter, S. Sarwate, M. L. Oelze, J. A. Zagzebski, M. N. Do, T. J. Hall, and W. D. O'Brien, "Analysis of human fibroadenomas using three-dimensional impedance maps," *IEEE Trans. Med. Imaging*, vol. 30, no. 6, pp. 1206–1213, 2011. (Cited on page(s): 45)

[135] M. I. Daoud and J. Lacefield, "Stochastic modeling of normal and tumor tissue microstructure for high-frequency ultrasound imaging simulations," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 12, pp. 2806–2815, 2009. (Cited on page(s): 45)

[136] M. Gyöngy, L. Balogh, K. Szalai, and I. Kalló, "Histology-based simulations of ultrasound imaging: Methodology," *Ultrasound Med. Biol.*, vol. 39, no. 10, pp. 1925–1929, 2013. (Cited on page(s): 45)

[137] H. Gao, H. F. Choi, P. Claus, S. Boonen, S. Jaecques, G. H. V. Lenthe, G. V. der Perre, W. Lauriks, and J. D'hooge, "A fast convolution-based methodology to simulate 2-D/3-D cardiac ultrasound images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, no. 2, pp. 404–409, 2009. (Cited on page(s): 45, 47)

DOI:10.15774/PPKE.ITK.2020.007

[138] B. Bürger, S. Bettinghausen, M. Radle, and J. Hesser, "Real-time GPU-based ultrasound simulation using deformable mesh models," *IEEE Trans. Med. Imaging*, vol. 32, no. 2, pp. 609–618, 2013. (Cited on page(s): 45)

[139] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Med. Biol. Eng. Comput.*, vol. 34, no. 1, pt. 1, pp. 1153–1163, 1996. (Cited on page(s): 47, 53, 77)

[140] X. Yan and X. G. Su, *Linear Regression Analysis: Theory and Computing*. Singapore: World Scientific, 2009. (Cited on page(s): 48)

[141] S. Weisberg, *Applied Linear Regression*, 3rd ed. Hoboken, NJ, USA:John Wiley & Sons, 2005. (Cited on page(s): 48)

[142] Onda Corporation. Tables of acoustic properties of materials. [Online]. Available: http://www.ondacorp.com/tecref_acoustictable.shtml (Cited on page(s): 50)

[143] E. Boni, L. Bassi, A. Dallai, F. Guidi, A. Ramalli, S. Ricci, J. Housden, and P. Tortoli, "A reconfigurable and programmable FPGA-based system for non-standard ultrasound methods," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 59, no. 7, pp. 1378–1385, 2012. (Cited on page(s): 50)

[144] V. D. Grosso and C. Mader, "Speed of sound in pure water," *J. Acoust. Soc. Am.*, vol. 52, no. 5B, pp. 1442–1446, 1972. (Cited on page(s): 51)

[145] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern*, vol. 9, no. 1, pp. 62–66, 1979. (Cited on page(s): 51)

[146] B. Luong. FFT-based convolution. Matlab Central File Exchange #24504. Last visited: 2019-12-11. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/24504-fft-based-convolution (Cited on page(s): 54)

[147] P. L. van Neer, M. G. Danilouchkine, M. D. Verweij, L. Demi, M. M. Voormolen, A. F. van der Steen, and N. de Jong, "Comparison of fundamental, second harmonic, and superharmonic imaging: A simulation study," *J. Acoust. Soc. Am.*, vol. 130, no. 5, pp. 3148–3157, 2011. (Cited on page(s): 62)

DOI:10.15774/PPKE.ITK.2020.007

[148] C. Marti, F. Varray, and C. Cachard, "A nonlinear ultrasound propagation simulator using the slowly varying envelope approximation," in *2014 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*. IEEE, 2014, pp. 5140–5144. (Cited on page(s): 62)

[149] K. J. Parker, "Superresolution imaging of scatterers in ultrasound B-scan imaging," *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4680–4689, 2012. (Cited on page(s): 64)

[150] Y. M. Benane, R. Lavarello, D. Bujoreanu, C. Cachard, F. Varrais, A. S. Savoia, E. Franchescini, and O. Basset, "Ultrasound bandwidth enhancement through pulse compression using a CMUT probe," *IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 6-9 Sept. 2017. (Cited on page(s): 65, 75)

[151] Y. M. Benane, R. Lavarello, D. Bujoreanu, C. Cachard, F. Varray, J.-M. Escoffre, A. Novell, and O. Basset, "Ultrafast ultrasound imaging using a resolution and bandwidth enhancement technique," *IEEE International Ultrasonics Symposium (IUS)*, 6-9 Sept. 2017. (Cited on page(s): 65, 75)

[152] S. Chen and K. J. Parker, "Enhanced axial and lateral resolution using stabilized pulses," *Journal of Medical Imaging*, vol. 4, no. 2, pp. 027 001–1–027 001–11, 2017. (Cited on page(s): 65, 75)

[153] A. Briggs, G. A. D. Briggs, and O. Kolosov, *Acoustic Microscopy*. Oxford University Press, 2010, vol. 67. (Cited on page(s): 74)

[154] R. A. Lemons and C. F. Quate, "Acoustic microscope – scanning version," *Applied Physics Letters*, vol. 24, pp. 163–165, 1974. (Cited on page(s): 74)

[155] K. Raum, "Microelastic imaging of bone," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, pp. 1417–1429, 2008. (Cited on page(s): 74)

[156] C. M. W. Daft and G. A. D. Briggs, "Wideband acoustic microscopy of tissue," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 36, pp. 258–263, 1989. (Cited on page(s): 74)

DOI:10.15774/PPKE.ITK.2020.007

[157] J. Mamou and D. Rohrbach, "Image formation methods in quantitative acoustic microscopy," *IEEE International Conference on Acoustics, Speech and signal Processing (ICASSP)*, pp. 6259–6263, 2017. (Cited on page(s): 74)

[158] M. K. H. Malo, D. Rohrbach, H. Isaksson, J. Töyräs, T. S. Jurvelin, I. S. Tamminen, H. Kröger, and K. Raum, "Longitudinal elastic properties and porosity of cortical bone tissue vary with age in human proximal femur," *Bone*, vol. 53, pp. 451–458, 2013. (Cited on page(s): 74)

[159] D. Rohrbach, H. O. Lloyd, R. H. Silverman, and J. Mamou, "Fine-resolution maps of acoustic properties at 250 MHz of unstained fixed murine retinal layers," *The Journal of the Acoustical Society of America*, vol. 137, pp. EL381– EL387, 2015. (Cited on page(s): 74)

[160] D. Rohrbach, E. Saegusa-Beecroft, J. Machi, E. J. Feleppa, and J. Mamou, "500-MHz quantitative acoustic microscopy imaging of unstained fixed 6-μm thin sections from cancerous human lymph nodes," *IEEE International Ultrasonics Symposium (IUS 2016)*, 18-21 Sept. 2016. (Cited on page(s): 74)

[161] K. Yoshida, Z. Deng, K. Ito, J. Mamou, H. Maruyama, H. Hachiya, and T. Yamaguchi, "Speed of sound, attenuation, and acoustic impedance of hepatic lobule in diseased rat liver observed by scanning acoustic microscopy with 250 MHz," *J. Acoust. Soc. Am.*, vol. 140, p. 3187, 2016. (Cited on page(s): 74)

[162] D. Rohrbach, A. Jakob, H. O. Lloyd, S. H. Tretbar, R. H. Silverman, and J. Mamou, "A novel quantitative 500-MHz acoustic microscopy system for ophthalmologic tissues," *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 715–724, 2016. (Cited on page(s): 74)

[163] S. Irie, K. Inoue, K. Yoshida, J. Mamou, K. Kobayashi, H. Maruyama, and T. Yamaguchi, "Speed of sound in diseased liver observed by scanning acoustic microscopy with 80 MHz and 250 MHz," *J. Acoust. Soc. Am.*, vol. 139, pp. 512–519, 2016. (Cited on page(s): 74)

DOI:10.15774/PPKE.ITK.2020.007

[164] Y. Saijo, M. Tanaka, H. Okawai, H. Sasaki, S.-I. Nitta, and F. Dunn, "Ultrasonic tissue characterization of infarcted myocardium by scanning acoustic microscopy," *Ultrasound in Medicine and Biology*, vol. 23, pp. 77–85, 1997. (Cited on page(s): 74)

[165] Y. Saijo, "Multimodal ultrasound microscopy for biomedical imaging," *Proc. Meet. Acoust.*, vol. 19, p. 75010, 2013. (Cited on page(s): 74)

[166] N. Hozumi, R. Yamashita, C.-K. Lee, M. Nagao, K. Kobayashi, Y. Saijo, M. Tanaka, N. Tanaka, and S. Ohtsuki, "Time-frequency analysis for pulse driven ultrasonic microscopy for biological tissue characterization," *Ultrasonics*, vol. 42, pp. 717–722, 2004. (Cited on page(s): 74)

[167] K. Miura, H. Nasu, and S. Yamamoto, "Scanning acoustic microscopy for characterization of neoplastic and inflammatory lesions of lymph nodes," *Sci. Rep.*, vol. 3, 2013. (Cited on page(s): 74)

[168] R. E. Baddour, M. D. Sherar, J. W. Hunt, G. J. Czarnota, and M. C. Kolios, "High-frequency ultrasound scattering from microspheres and single cells," *Journal of Acoustical Society of America*, vol. 117, pp. 934–943, 2005. (Cited on page(s): 74)

[169] L. R. Taggart, R. E. Baddour, A. Giles, G. J. Czarnota, and M. C. Kolios, "Ultrasonic characterization of whole cells and isolated nuclei," *Ultrasound in Medicine & Biology*, vol. 33, pp. 389–401, 2007. (Cited on page(s): 74)

[170] E. C. Weiss, P. Anastasiadis, G. Pilarczyk, R. M. Lemor, and P. V. Zinin, "Mechanical properties of single cells by high-frequency time-resolved acoustic microscopy," *IEEE UFFC*, vol. 54, pp. 2257–2271, 2007. (Cited on page(s): 74, 76)

[171] J. A. Hildebrand, D. Rugar, R. N. Johnston, and C. F. Quate, "Acoustic microscopy of living cells," *Proceedings of the National Academy of Sciences*, vol. 78, pp. 1656–1660, 1981. (Cited on page(s): 74)

DOI:10.15774/PPKE.ITK.2020.007

[172] E. M. Strohm, G. J. Czarnota, and M. C. Kolios, "Quantitative measurements of apoptotic cell properties using acoustic microscopy," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, pp. 2293–2304, 2010. (Cited on page(s): 74)

[173] J. Litniewski and J. Bereiter-Hahn, "Measurements of cells in culture by scanning acoustic microscopy," *Journal of microscopy*, vol. 158, pp. 95–107, 1990. (Cited on page(s): 74)

[174] Y. Saijo, H. Sasaki, M. Sato, S. I. Nitta, and M. Tanaka, "Visualization of human umbilical vein endothelial cells by acoustic microscopy," *Ultrasonics*, vol. 38, pp. 396–399, 2000. (Cited on page(s): 74)

[175] F. Foroozan, M. A. O'Reilly, and K. Hynynen, "Microbubble localization for 3D super-resolution ultrasound imaging using curve fitting and deconvolution methods," *IEEE Transactions on Biomedical Engineering*, 2018. (Cited on page(s): 74)

[176] O. Michailovich and A. Tannenbaum, "Blind deconvolution of medical ultrasound images: a parametric inverse filtering approach," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 16, pp. 3005–3019, 2007. (Cited on page(s): 74, 75)

[177] A. Basarab, D. Rohrbach, N. Zhao, J.-Y. Tourneret, D. Kouamé, and J. Mamou, "Enhancement of 250-MHz quantitative acoustic-microscopy data using a single-image super-resolution method," *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pp. 827–830, 2017. (Cited on page(s): 74, 75, 77, 80)

[178] F. Sroubek, G. Cristóbal, and J. Flusser, "A unified approach to superresolution and multichannel blind deconvolution," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 16, pp. 2322–2332, 2007. (Cited on page(s): 74)

DOI:10.15774/PPKE.ITK.2020.007

[179] F. Sroubek, J. Flusser, and M. Sorel, "Superresolution and blind deconvolution of video," *19th International Conference on Pattern Recognition*, 8-11 Dec. 2008. (Cited on page(s): 74)

[180] X.-H. Han, J. Wang, B. Shi, Y. Q. Zheng, and Y.-W. Chen, "Hyper-spectral image super-resolution using non-negative spectral representation with data-guided sparsity," *IEEE International Symposium on Multimedia (ISM)*, 11-13 Dec. 2017. (Cited on page(s): 74)

[181] J. Guan, S. Luo, and D. Li, "Blind deconvolution-based angular superresolution algorithm for scanning radar system," *CIE International Conference on Radar (RADAR)*, 10-13 Oct. 2016. (Cited on page(s): 74)

[182] F. A. Dharejo, Z. Hao, A. Bhatti, M. N. Bhatti, J. Ahmed, and M. A. Jatoi, "Improved dictionary learning algorithm with mappings for single image super-resolution," *IEEE International Conference on Imaging Systems and Techniques (IST)*, 18-20 Oct. 2017. (Cited on page(s): 74)

[183] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: a review," *IEEE Signal Processing Magazine*, vol. 34, pp. 85–95, 2017. (Cited on page(s): 74)

[184] J. Hatvani, A. Horváth, J. Michetti, A. Basarab, D. Kouamé, and M. Gyöngy, "Deep learning-based super-resolution applied to dental computed tomography," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, pp. 120–128, 2018. (Cited on page(s): 74, 79, 80, 91)

[185] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," *Proc. 13th European Conf. Computer Vision (ECCV)*, pp. 184–199, 2014. (Cited on page(s): 74, 91)

[186] Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, and H. Wang, "Deep learning microscopy," *Optica*, vol. 4, pp. 1437–1443, 2017. (Cited on page(s): 74, 91)

DOI:10.15774/PPKE.ITK.2020.007

[187] T. Taxt and G. V. Frolova, "Noise robust one-dimensional blind deconvolution of medical ultrasound images," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 46, pp. 291–299, 1999. (Cited on page(s): 75)

[188] T. Taxt, "Three-dimensional blind deconvolution of ultrasound images," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 48, pp. 867–871, 2001. (Cited on page(s): 75)

[189] Z. Chen, A. Basarab, and D. Kouamé, "Compressive deconvolution in medical ultrasound imaging," *IEEE transactions on medical imaging*, vol. 35, pp. 728–737, 2016. (Cited on page(s): 75)

[190] J. H. Kim, P. R. Hill, N. Canagarajah, D. Rohrbach, D. Kouamé, J. Mamou, A. Achim, and A. Basarab, "Compressed quantitative acoustic microscopy," *IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2017. (Cited on page(s): 75)

[191] J. Reeg and M. L. Oelze, "Improving lateral resolution in ultrasonic imaging by utilizing nulls in the beam pattern," *IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2015. (Cited on page(s): 75)

[192] L. Roquette, M. Simeoni, P. Hurley, and A. Besson, "On an analytical, spatially-varying, point-spread-function," *IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 6-9 Sept. 2017. (Cited on page(s): 75)

[193] Z. Chen, A. G. J. Besson, J. P. Thiran, and Y. Wiaux, "A novel concept of deconvolution for ultrasound imaging," *International BASP Frontiers workshop*, vol. EPFL-CONF-223426, 2017. (Cited on page(s): 75)

[194] T. Taxt and R. Jirik, "Superresolution of ultrasound images using the first and second harmonic signal," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 51, pp. 163–175, 2004. (Cited on page(s): 75)

[195] T. Taxt and J. Strand, "Two-dimensional noise-robust blind deconvolution of ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 48, pp. 861–866, 2001. (Cited on page(s): 75)

DOI:10.15774/PPKE.ITK.2020.007

[196] X. Wang, C. J. Ritchie, and Y. Kim, "Elevation direction deconvolution in three-dimensional ultrasound imaging," *IEEE transactions on medical imaging*, vol. 15, pp. 389–394, 1996. (Cited on page(s): 75)

[197] R. Mohammadi and A. Mahloojifar, "Resolution improvement of scanning acoustic microscopy using sparse signal representation," *Journal of Signal Processing Systems*, vol. 54, pp. 15–24, 2009. (Cited on page(s): 75)

[198] S. H. Chan, K. Ramsin, G. B. Kristofor, G. E. Philip, and N. Q. Truong, "An augmented Lagrangian method for total variation video restoration," *IEEE Transactions on Image Processing*, vol. 20, pp. 3097–3111, 2011. (Cited on page(s): 77)

[199] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, ..., and M. Kudlur, "Tensorflow: A system for large-scale machine learning," *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283, 2016. (Cited on page(s): 79)

[200] F. J. Huang, Y. L. Boureau, and Y. LeCun, "Unsupervised learning of invariant feature hierarchies with applications to object recognition," *IEEE conference on computer vision and pattern recognition*, pp. 1–8, 2007. (Cited on page(s): 79)

[201] O. Ronneberger, F. Philipp, and B. Thomas, "U-Net: Convolutional networks for biomedical image segmentation," *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, 2015. (Cited on page(s): 79)

[202] P. J. Huber, "The place of the l1-norm in robust estimation," *Computational Statistics & Data Analysis*, vol. 5, pp. 255–262, 1987. (Cited on page(s): 79)

[203] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, pp. 1285–1298, 2016. (Cited on page(s): 89)

DOI:10.15774/PPKE.ITK.2020.007

[204] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015. (Cited on page(s): 90)

[205] J. Li, F. Fang, K. Mei, and G. Zhang, "Multi-scale residual network for image super-resolution," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 517–532, 2018. (Cited on page(s): 90)

[206] J. Mamou, T. Pellegrini, D. Kouame, and A. Basarab, "A convolutional neural network for 250-MHz quantitative acoustic-microscopy resolution enhancement," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 6212–6215. (Cited on page(s): 91)

DOI:10.15774/PPKE.ITK.2020.007

Acknowledgements

I can still remember the day back in 2012 when I was looking for a topic and a supervisor for my Project Laboratory. I skipped class to visit Dr. Miklós Gyöngy in his office to talk about one of the topics. When he was explaining what the research would be about, he mentioned something and I was not sure that it was entirely correct, so I formulated my doubts in a question. After a short silence, his answer stuck with me: "I am not entirely sure so I have to do some re-cap on this". I think at that moment I realized I am at the right place. He was (and still is) not afraid to admit if he does not know something, but is ready to investigate it or make a plan on how to tackle the problem. I am really happy because he was not just my supervisor, but also became my friend during our time working together. This thesis would not have been possible without his dedicated effort. I really learnt a lot from him not just professionally but in human qualities as well, and I would like to say thank you.

I would also like to thank Dr. Louis Atallah and Dr. Hans Weda for being my supervisors during my ERASMUS internship at Philips (Eindhoven, the Netherlands) in 2014. I am totally aware that back then my spoken English was not the best (let's admit it was quite terrible), yet they were very patient and understanding with me, and treated me as an equal colleague of theirs. I can still remember I was literally sweating when trying my best to talk to people in English. Even though it was before my PhD studies, I think the whole internship was a big milestone for me and had a big impact on me. That time I realized that I should not be afraid if I am not an expert in something, but I am willing to work for it. I will always be grateful to them for providing this opportunity.

I would like to express my gratitude to Dr. Wolfgang Bost (Fraunhofer IBMT, Sulzbach), who sacrificed a lot of his time in order to help me in the research we have done together. I am also indebted to Dr. Marc Fournelle (Fraunhofer IBMT,

DOI:10.15774/PPKE.ITK.2020.007

Sulzbach) who accepted me for an ERASMUS internship so that I was able to work on my PhD research with one of the best scanning acoustic microscope systems available.

I am particularly grateful to Janka Hatvani, who always found time in the last 8 years to help me if I was stuck or did not understanding something, whether it was regarding math, coding, deep learning, physics or ultrasound. Despite we are both PhD students and about the same age, I feel like she is my second supervisor. I am sure without her help and friendship I would not have been able to finish my Bachelor's degree, yet alone my PhD. She was also kind enough to proofread most of my dissertation and thesis booklets in order to improve the quality of the manuscript.

I am also grateful to Dr. András Horváth and Dr. Imre Kalló who were open to do a research together, even with their limited amount of time.

I would like to thank my labmates, Gergely Csány, Krisztián Füzesi, Péter Marosán, for their encouragement and fun times while working together in the last couple of years.

I would also like to acknowledge the time of the reviewers spent on reviewing the manuscript. Their constructive and detailed critics helped me to considerably improve the dissertation.

I gratefully acknowledge the financial support of the NTP-NFTÖ-17 Nation's Young Talent Scholarship, the ÚNKP-18-3 and ÚNKP-19-3 New National Excellence Program, all through the Hungarian Ministry for Innovation and Technology. I am also extremely grateful for the financial support of the Pázmány Péter Catholic University through the KAP R&D and equipment Grants (2014-2018) and EFOP-3.6.2-16-2017-00013. This research also has been partially supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.3-VEKOP-16-2017-00002).

Last but not least, I would like to thank my parents, Judit Makráné Biczák and István Makra, for providing me the best education possible so that I was able to pursue a PhD degree, and being supportive of all my decisions. They have done inhuman efforts in order to be able to raise and support me and my siblings, and

DOI:10.15774/PPKE.ITK.2020.007

help us in every way possible, even to this day. I would like to dedicate this thesis to them.