Design of Imaging Systems for Depth Sensitivity and Detection of Object Growth

by

Andrew Marc Stein

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Masters of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2002
(June 2002)

Author ....
Department of Mechanical Engineering
May 24, 2002

Certified by ....
George Barbastathis
Assistant Professor
Thesis Supervisor

Accepted by ....
Ain A. Sonin
Chairman, Department Committee on Graduate Students

Design of Imaging Systems for Depth Sensitivity and Detection of Object Growth

by

Andrew Marc Stein

Submitted to the Department of Mechanical Engineering
on May 24, 2002, in partial fulfillment of the
requirements for the degree of
Masters of Science

Abstract

This thesis is divided into two parts. Part I assesses the feasibility of an ultrasound diagnostic tool for taking monthly breast measurements in order to detect breast cancer sooner than it can be detected in a mammogram. It describes simple phantom experiments in which a 0.3 cm^3 mass could be detected with 100% accuracy. A finite element breast model produced similar results. Part II describes a tradeoff between shift invariance and depth resolution that seems to occur in almost all imaging systems. A metric for shift variance and depth resolution are described and are then applied to three different imaging systems.

Thesis Supervisor: George Barbastathis
Title: Assistant Professor

Acknowledgments

I would first like to acknowledge my Advisor, George Barbastatathis. He has taught me a great deal and allowed me the freedom to investigate a wide variety of interests. He has also been patient with me in my search for a topic that I truly find interesting. I could not have learned as much as I did here without his help. I would also like to thank the graduate students with whom I shared my lab, Wei-Chuan Shih, Troy Savioe, Greg Nielson and Arnab Sinha. Their companionship throughout these past two years has made my work here much more enjoyable.

Professor Alex Slocum, Xueen Yang, Asha Balakrishnan, and Kateri Garcia were great coresearchers. The phantom experiments could not have been done without Asha and Kateri's injection device. Xueen helped a lot with the experiments and with writing our NIH grant proposal. She also gave me some terrific maple butter which I am still enjoying at the time of my thesis completion. And Alex's advice, critiques, guidance, and excitement were great assets to this project. Dr. Reuben Mezrich was the person responsible for giving us our initial funding, and he provided constant help by finding us an ultrasound lab to work in and finding angioplasty balloons from Boston Scientific for us to use in our experiments.

Professor Kullervo Hynynen and his lab group, including Greg, Sham, Tonya, Shuba, Kristina, Jason, Randy and José, were a tremendous amount of help when running the ultrasound experiments. Their advice and guidance made the ultrasound experiments possible and they were always happy to answer all of my questions, even the ones that I really should have known the answer to. In addition, they are the friendliest lab group I have ever worked with and working there was always fun and interesting. Weidlinger Associates and Dave Vaughn provided the Finite Element software for this research. Dave was a tremendous help. He quickly answered all questions I had about their software and he also gave much helpful advice for running simulations.

I would also like to thank Professors Sanjay Sarma, John Brisson, Terry Orlando. The high pressure environment at MIT forces many professors to focus primarily on their research while teaching typically takes a back seat. I was fortunate enough during my time here, to take 2.007, 2.008, and 2.998 with Professor Sarma, 2.005 and 2.006 with Professor Brisson, and 6.728 with Professor Orlando. It was apparent in all of these classes that these professors put a great deal of effort into their courses and because of this, I learned a lot, both in lecture and in office hours where they were always happy to answer lots of questions.

Fred Jao, Ian Ross, Mike Perry, Mike Salamina, Eric Ferreira, Mike Dewberry Gavin Sy, Mealani Nakamura, Dan Amano, and Rusty Lee have all been good friends of mine for at least the past 5 years. They have just made life way more interesting and livable and I think I've probably learned as much or more from them than from my classes and research. I'm sure at least half of them would probably make fun of me for the sincerity of this acknowledgements section, but I do mean everything I've said here and the acknowledgements section is really no place for bitterness, cynicism or wit.

Finally, I would like to thank my parents. They have always wanted the best for me and they have worked hard for many years to provide me with opportunities that they never had. They provided a loving and nurturing environment for me to grow up in and even after I moved away, they continued their support. I certainly could not have made it here without them.

This work was supported by CIMIT, Rosenblith Graduate Fellowship, and NSF.

Contents

List of Figures

List of Tables

Part I

Ultrasound Detection of Tumor Growth

Chapter 1

Ultrasound Introduction

Currently, 1 in 8 women in the United States are expected to develop breast cancer [4]. The main form of imaging used to detect breast cancer is the mammogram, which is recommended approximately once a year once a woman reaches 40. Due to low contrast and noise, mammograms, miss roughly 10% of all tumors [5]. One could improve tumor visibility by comparing images taken at different times, but due to the non-rigid nature of tissue, the breast is never positioned exactly the same with each measurement and thus the two mammograms will not be aligned. One solution to this problem is to develop elaborate registration algorithms to align multiple images. Many people are working on this but image registration is still a major unsolved problem in digital mammography [6]. Another possible solution is to take multiple mammograms at each sitting to ensure that at least two will align, but because x-rays in high doses cause cancer, the number of images that can be taken per year is limited.

An alternative form of imaging that avoids ionizing radiation is ultrasound. With ultrasound, many measurements of the breast can safely be taken and thus measurements taken at different times could be more easily aligned without the need for complex image registration algorithms. While ultrasound has significantly worse resolution than mammography, forming high quality images is not necessarily a requirement for detecting tumors. If measurements can be accurately aligned and compared and if changes in the healthy tissue are small compared to tumor growth, a high quality image will not be necessary for detection. An additional improvement on mammography is that with ultrasound, measurements can be taken monthly, instead of yearly, thus further increasing the probability of detection.

Our goal was to investigate the feasibility and performance of an algorithm that obtained monthly ultrasound data from precisely aligned transducers. Accomplishing this task required two innovations. The first was in determining how to generate growth data that would resemble biological data. To address this task, we ran phantom experiments and compared these experiments to a Finite Element simulation. The second innovation was that once realistic data was created, we needed to develop a detection algorithm. The major detection algorithm used had two major parts, data reduction, using Principal Component Analysis, and classification, using Nearest Neighbor Learning. We will conclude with plans for future work.

Chapter 2

Fundamental Physics

2.1 Ultrasound Propagation in Tissue

Waves propagate through air by the compression and expansion of the air molecules. These waves are known as acoustic waves. At frequencies (ν) between 20 Hz and 20 kHz, they are within the range of human hearing and are called sound waves. Similar physical principles show that these compression waves also propagate through other media, including water, steel, and biological tissue. Ultrasound is a simply a compression wave with a frequency that is higher than 20 kHz. Many linear waves that can be found in nature, including light, sound, and vibration. Simple wave propagation can be described by Equation 2.1, where x = position, t = time, c = the wave propagation speed, and ξ can be any wave variable including pressure (p), velocity (u), density (ρ) etc.

$$\frac{\partial^2 \xi}{\partial t^2} = c^2 \frac{\partial^2 \xi}{\partial x^2} \tag{2.1}$$

Solutions to this equation are of the form $f(x \pm ct)$, and the sinusoidal solution to this problem is $\xi = \xi_0 e^{ik(x \pm ct)}$, where k = the wave number. Other important fundamental equations to be aware of are $c = \lambda\nu$, where λ = wavelength, $k = \nu/2\pi$, and $\omega = 2\pi\nu$, where ω is the angular frequency, and $c = \sqrt{K/\rho}$, where K is the bulk modulus of the medium and ρ is the density. Some intuition for this last equation can

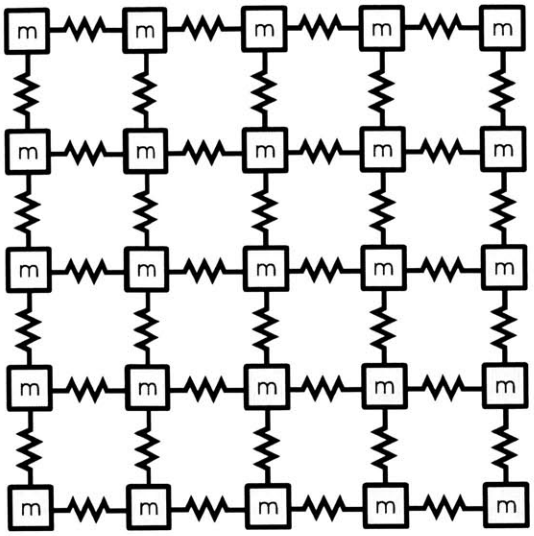


Figure 2-1: A system made of identical masses and springs with dynamics that follow the wave equation as the spacing between elements becomes infinitely dense

be gained by imagining the propagation medium as a meshwork of uniform springs and masses as shown in Figure 2-1. We know that by solving the simple, homogeneous ODE $m\ddot{x} + kx = 0$ for a spring-mass system with no damping, the natural frequency equals $\sqrt{k/m}$. The higher the natural frequency, the sooner a displacement of an individual element will affect its neighboring elements, leading to a faster disturbance propagation and a higher c. For the continuous medium, we simply have ρ instead of m and K instead of k. An overview of acoustics with a detailed derivation of the wave equation can be found in [7] and a brief overview can be found in [8].

So clearly, when modelling wave propagation in tissue, one must know c. If the propagation medium is perfectly homogeneous and lossless, then c is the only parameter we need. The next level of complexity is when a plane wave propagates through a homogeneous medium and encounters an interface with a second homogeneous medium as shown in Figure 2-2. Like light waves, acoustic waves obey the laws of reflection and refraction. $\theta_r = \theta_i$ and Snell's law, $\frac{\sin\theta_i}{\sin\theta_t} = \frac{c_1}{c_2}$ gives θ_t. To determine how much of the wave is transmitted to the second medium, one must calculate the reflection and transmission coefficients, as described in Chapter 6 of [7]. To do so, one needs to know, the impedance of each medium (Z). For a plane wave, $Z = \rho c$.

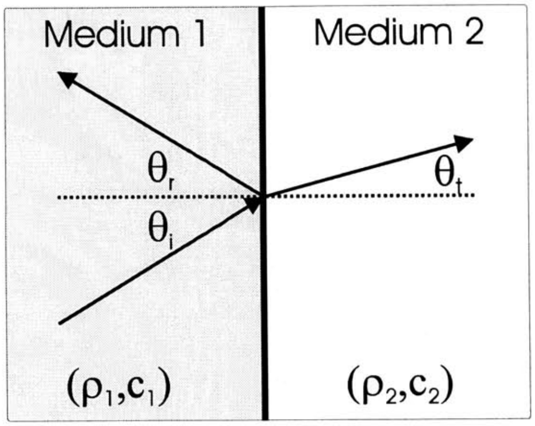


Figure 2-2: The reflection and refraction of a plane wave at an interface

Thus when multiple media are present, one must also know ρ.

Parameters c and ρ are relatively easy quantities to measure and can be found in [2] and [3] respectively. While tissue is known to support shear waves, these waves are damped locally, and their energy dissipates. It turns out that their dissipation can be included with the attenuation [9].

Another level of complexity appears when we account for the fact that as a wave propagates through a medium, even if it is homogeneous, its energy dissipates as it propagates. The damped wave is expressed as $\xi = \xi_0 e^{-\alpha x} e^{ik(x \pm ct)}$, where α is the bulk attenuation coefficient. This dissipation can be divided into three categories: viscous losses, heat losses, and molecular losses. Viscous losses result from relative motion between different parts of the medium. Heat losses occur because regions of the fluid that are under a higher pressure are also at a higher temperature; some of this heat flows to the lower pressure and temperature regions through conduction. Molecular losses occur when some of the wave energy is transferred to vibrational energy of the molecules within the medium. More discussion can be found in Chapter 9 of [7].

Additional losses come from the wave scattering off small particles inside the fluid. Scattering in soft tissue is a very complex phenomenon that is not easy to quantify. A good overview can be found in [8]. Analytical solutions for plane wave scattering off solid cylinders and spheres was derived by Faran [10] and elliptical geometries have also been solved for [11]. However, because scattering in general is very difficult

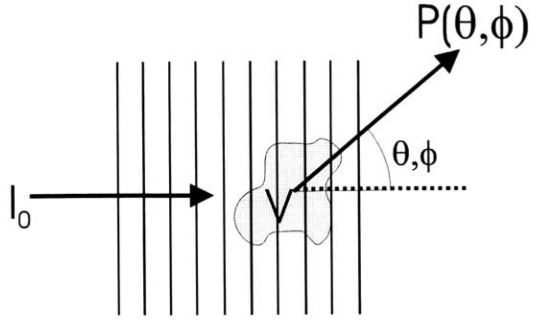


Figure 2-3: Representation of backscattering coefficient. I_0 is the intensity of the incident plane wave, V is the volume of the scatterer and P is the scattered power.

to model, it is often modelled by two bulk parameters; the backscatter coefficient (σ) and the differential backscatter coefficient (σ_d). σ_d is a measure of the power scattered in a given direction by a unit volume of a material that is hit with a plane wave of unit intensity as shown in Figure 2-3. It is defined according to Equation 2.2, where $V =$ the volume of the material, I_0 is the intensity of an incoming plane wave, and $P_\Omega(\theta, \phi)$ is the scattered power per solid angle as a function of the angles θ and ϕ. $\sigma = \iint \sigma_d d\theta d\phi$. To eliminate confusion, it should be noted that σ has units of $1/(mSr)$. $I = \frac{1}{2}pu$ which has units of W/m^2 while P_Ω has units of W/Sr.

$$\sigma = \frac{1}{V}\frac{P_\Omega(\theta, \phi)}{I_0} \tag{2.2}$$

The material properties of actual biological tissue are shown in Table 2.1. It should be noted scattering measurements made by different labs can vary by a factor of 10, as can be seen when comparing [12], [1], [13], and [14]. Moreover, in [15], phantoms with identical properties were made in Madsen's laboratories and then sent to 10 different groups for measurement. Even among these measurements, σ varied by a factor of 10. While this variance is discouraging, the σ ratio between various tissue types for a given measurement technique has much higher agreement across labs.

It should be noted that a fifth property of tissue, its nonlinearity, often referred to as B/A, has not been discussed at all. With low amplitude waves, nonlinear effects become negligible, but when the pressure rises, the wave propagation can change dramatically. An overview of nonlinear ultrasound can be found in [16]. In our

Material	ρ (kg/m^3)	c (m/s)	α_0 (dB/cm)	n_α -	σ_0 (10^{-3} (cm Sr MHz)$^{-1}$)	n_σ -
Water	1000	1488	0	-	0	-
Fat	937	1440	.158	1.7	9.5×10^{-5}	1.9
Parenchyma	1060	1540	.87	1.5	5.2×10^{-3}	1.3
Tumor	1080	1580	.57	1.3	2.6×10^{-4}	0.8

Table 2.1: Tissue Properties. $\sigma(f) = \sigma_0 f^n$ and $\alpha(f) = \alpha_0 f^n$, where f is the frequency in MHz. The attenuation and scattering properties were taken from [1]. The c values were estimated from figures in [2]. The ρ values were taken from [3]

analysis, we will be assuming small displacements and we ignore nonlinear effects.

2.2 Piezoelectric Transducers

Ultrasound waves are typically generated using piezoelectric transducers. Piezoelectric materials have the property that when a voltage is placed across them, they deform, and when a force is exerted on them, they create a voltage drop. Piezoelectric materials exhibit this behavior because they are made up of polar, non-sperical molecules, as shown in Figure 2-4. At rest, the molecules all have a random orientation. When a voltage is applied across the materials, the molecules align to minimize their energy, causing the material to expand in one dimension and contract in another. If the voltage applied to the material is a 5MHz sinusoid, then the transducer will create a 5MHz acoustic wave. Conversely, if a force is exerted on the piezoelectric material, the molecules inside will also align, causing a voltage drop to be present which can be measured. Thus the same transducer that sends a pulse can receive it as well. Common piezoelectric material used in ultrasound transducers are lead zirconate titanate (PZT) and polyvinylidene fluoride (PVDF). A more in depth discussion of piezoelectric materials can be found in [17].

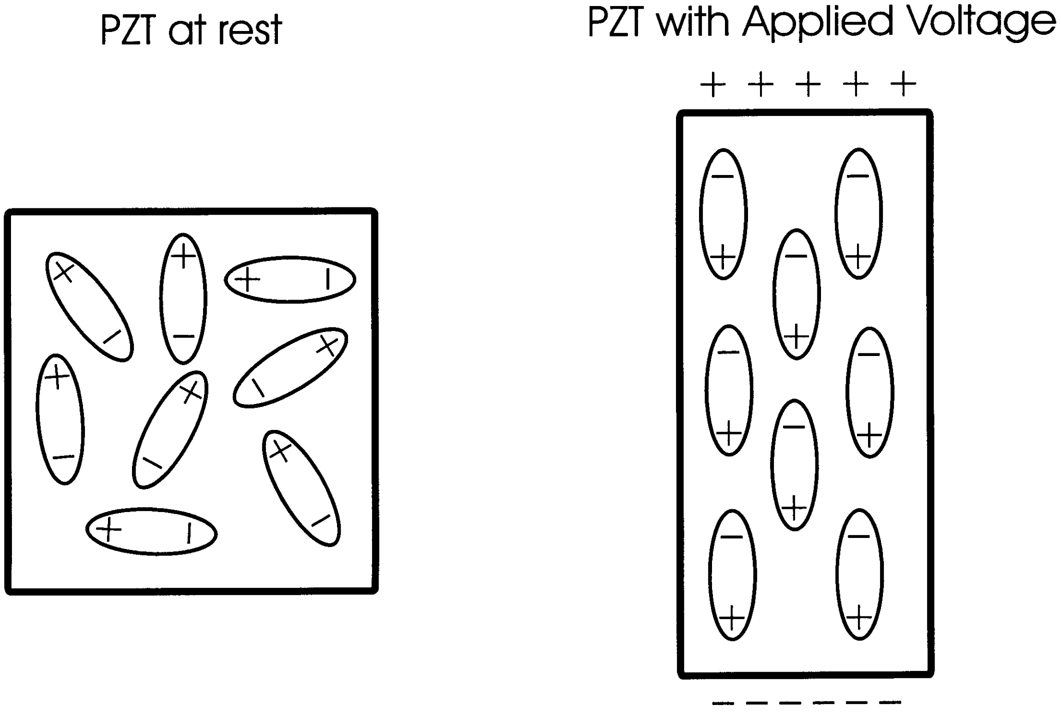


Figure 2-4: PZT at rest and with an applied voltage or displacement

Chapter 3

Phantom Experiments

To test the feasibility of a differential detection algorithm and get a feel for what a growing tumor might look like, tissue mimicking (TM) breast phantoms were created. Ernest L. Madsen, at University of Wisconsin developed TM material for the purpose of testing ultrasound equipment and many aspects of our phantom design came from his work [18], [19], and [20]. As discussed in Chapter 2, when creating TM phantoms that have the acoustic properties of biological tissue, there are four important bulk properties to match: c, ρ, α and σ.

This thesis will discuss two types of phantoms that were developed. The first used gelatin to simulate breast tissue and vegetable shortening to simulate a tumor. The second involved filling angioplasty balloons with solutions of water and glass beads to simulate tumor and glandular tissue. The gelatin phantoms were the only ones to yield useful and qualitatively reasonable results and so they will be discussed first. A discussion as to why the angioplasty balloons failed will be saved for Chapter 5

3.1 Experimental Setup

In our first round experiments, we chose to construct a very simple phantom made of gelatin and water. There are two major materials that can be used for constructing phantoms: gelatin and agar. These materials have similar speeds of sound and densities to tissue, but in their pure form, they have very low attenuation and back scatter.

Gelatin consists mostly of a collagen, a major structural protein that is present in bone, hooves, and cartilage. Agar is a structural galactose polymer found in the cell walls of red algae. Both come in powder and they create gels by linking together and forming large molecular networks when placed in water. Advantages of agar are that it has a high melting point of 78 C and it is resistant to bacterial decay, whereas gelatin melts at room temperature and decays within weeks. On the other hand, gelatin mixes more uniformly than agar and since we were not concerned with long term phantom storage at room temperature, we chose gelatin for our initial experiments. To preserve the gelatin phantom, one can add formaldehyde, as described in [20], but this also was unnecessary for our experiments.

Our first set of phantoms were created by mixing one knox gelatin packet (7.1 g) with one cup of water inside of a hemispherical mold with a 4.6 cm inner radius. These phantoms were placed in a cold storage room at 6 C overnight or longer.

We used the experimental setup shown in Figure 3-1 to measure the speed of sound through the gelatin. We placed our phantom on top of a piece of metal and then placed a transducer on top of the hemisphere such that it was parallel to the metal base. The transducer sent a pulse through the gelatin which reflected off the stage and returned to the transducer. The pulse continued to bounce back and forth between the transducer and metal baseplate. A plot of the received signal is shown in Figure 3-2. Then, by using the simple relation $c = d/t$ where d = distance the pulse traversed (9.2 cm) and t = time of flight (60 μs), corresponding to distance between peaks in Figure 3-2, we could easily estimate c to be 1530 m/s.

Since our phantom is composed almost completely of water, we know that its density is very close to 1000 kg/m^3. In these phantoms, backscatter and attenuation are very low. We tried mixing the gelatin with graphite, oil, and glass beads as described in in [18] and [20]. Unfortunately, we did not have a device with which to rotate the gel as it set and thus the graphite settled to the bottom while the oil floated to the top. Rather than spend more time developing more complex phantoms, we chose to continue using gelatin. The scattering and attenuation are added in the postprocessing stage.

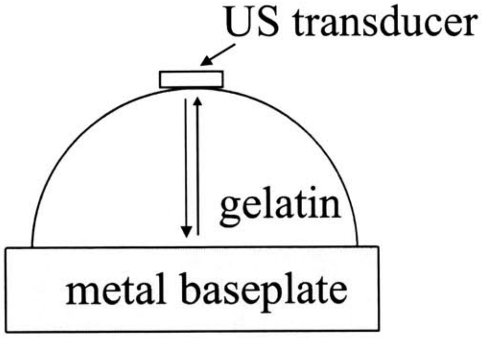


Figure 3-1: Experimental Setup for measuring the speed of sound through gelatin

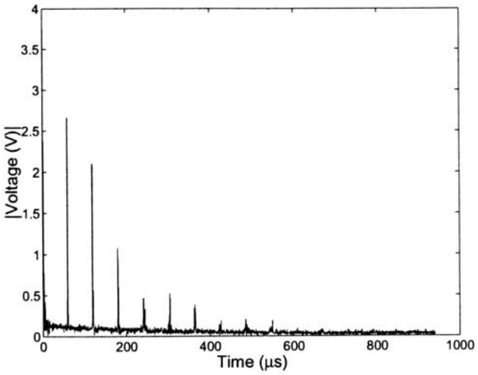


Figure 3-2: Data taken from speed of sound experiments. Each peak corresponds to the detection of a reflected pulse.

To simulate tumor growth, we injected vegetable shortening into the phantom through a syringe. Vegetable shortening is essentially fat and so it has very similar acoustic properties. The shortening was injected at a rate of approximately 2 mm^3/s and data was recorded from the transducers every 0.4 s. It should be noted that the vegetable shortening growth was not predictable. The "tumors" maintained a disk shape as they grew. At a full size of 0.3 cm^3, the disk had a diameter of approximately 15 mm and a thickness of 1.5 mm. The orientation of the disk, which can be defined as the direction in which the normal to its surface pointed, was unpredictable and seemed random. We believe that the disk orientation is determined by the fracture plane that develops in the gelatin when the needle is inserted. In addition, the position of the tumor with respect to the needle was not controllable and could vary by as much as 15 mm. Since the 6 dB beam focus of our transducer was 1.6 mm, this variation could lead to experiments in which the ultrasound pulse completely missed the tumor and nothing was detected. This problem could have been avoided if we had the capability to acquire data from multiple transducers at once.

A phantom stage was designed by Kateri Garcia for positioning the syringe and the phantom. The phantom stage is shown in Figure 3-3. The syringe is fit into a syringe holder made of aluminum, which is attached to a lead screw for horizontal translation, bringing the needle into and out of the breast phantom. The distance travelled by

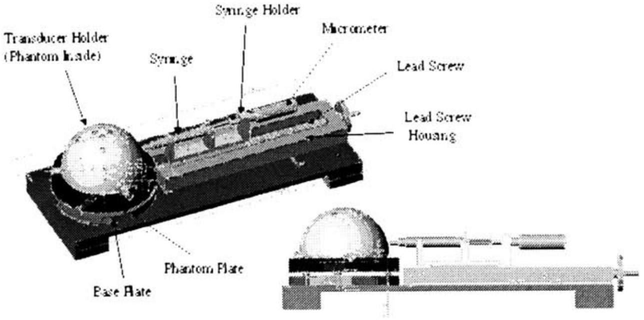


Figure 3-3: AutoCAD rendering of the phantom stage showing the isometric and side views. Designed by Kateri Garcia.

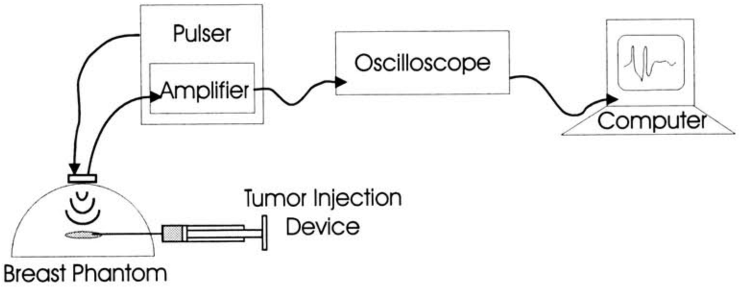


Figure 3-4: Schematic of data acquisition apparatus.

the syringe can be read from an embedded scale on the lead screw housing. When the syringe is positioned, a micrometer attached to the syringe pump is used to control the volume injection. The phantom is placed on a teflon plate, which is positioned on the base plate by a kinematic coupling which allows it to be repeatably located should it be desired to move the phantom. The transducer holder, which is a plastic shell that has the same size as the phantom, allows for the placement of cylindrical transducers.

A simple schematic of the test apparatus is shown in Figure 3-4. We use a 5 MHz, single element contact type ultrasound transducer from *Panametrics*. Table 3.1 lists its characteristics. Signal generation, 59 dB of amplification, and bandpass filtering between 1-10 MHz is done by a *Panametrics Model* 5072PR pulser/receiver, which can

Freq.	5 MHz
Diameter	0.25 in.
Near Field Distance	1.309 in.
Focal Length Point Target	0.43 in.
Beam Diameter (-6 dB)	0.0642 in.
Focal Zone	1.7877 in.

Table 3.1: Parameters of the single element ultrasound transducer used in our experiments (5 MHz). All parameters are measured in water.

be operated in both transmit/receive and transmission mode. The data is displayed on an oscilloscope, which is connected to a GPIB board and a PC for data acquisition. The sampling rate was 50 MS/s.

It should be noted that we used contact transducers when we should have used immersion transducers. Contact transducers have a high impedance for matching other high impedance solids like aluminum and steel. Immersion transducers are designed to go under water and thus they have a low impedance so that they couple well to media like gelatin and water. It would have made more sense to use immersion transducers. By using high Z transducers, much of the input pulse is reflected straight back off the transducer/coupling gel interface, resulting in a weaker signal.

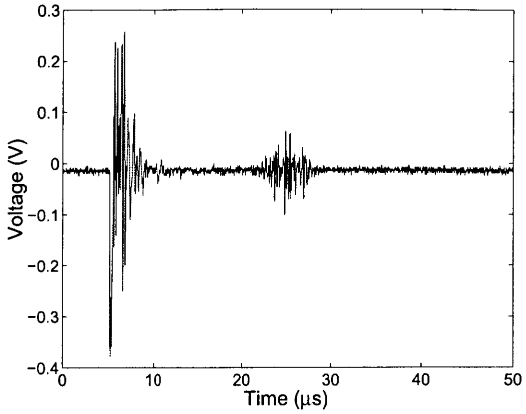


Figure 3-5: Raw Data from a single transducer

3.2 Phantom Experiment Data Analysis

3.2.1 The Preprocessing Stage

Some sample data taken from the transducer are shown in Figure 3-5. There are two major features that occur in this data. One at 5 μs and one at 25 μs. The 5 μs feature comes from the electrical signal sent to the transducer from the pulser. Since the electrical impedance of the cable does not exactly equal the electrical impedance of the transducer, some of this signal gets reflected back to the oscilloscope. While this feature of the data occurs at the same time of the initial pulse, it does not represent what this pulse actually looks like and is ignored in any future data processing.

The second feature at 25 μs comes from reflections off the vegetable shortening. If the shortening were perfectly homogeneous and its boundaries with the gelatin were smooth, one would expect to see a few perfect reflections from the gelatin-shortening interface. Even if the boundaries were rough, we would expect a more distorted signal, but it would have a more uniform envelope rather than the triangular envelope that we see. The reason we see this triangular shape is due to backscatter from the vegetable shortening. The vegetable shortening is nonhomogeneous and there are likely to be small air bubbles inside that scatter the acoustic wave as it passes through. The scattered power should increase linearly with volume as described in Chapter 2. The actual shape of the tumor is typically narrow at the edges and wide at the center,

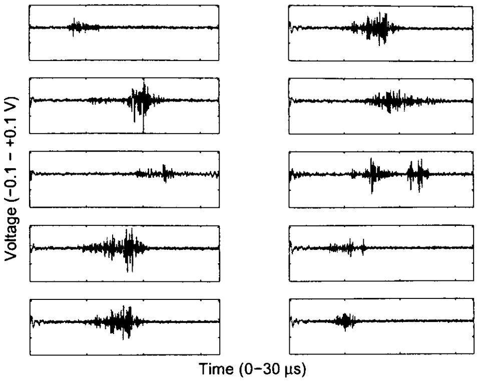


Figure 3-6: The waveform of 10 tumors that have reached their final volume of 0.3 cm^3.

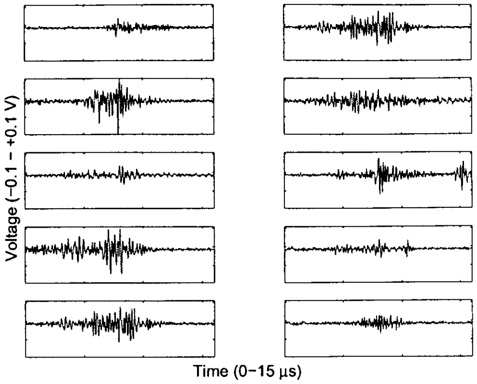


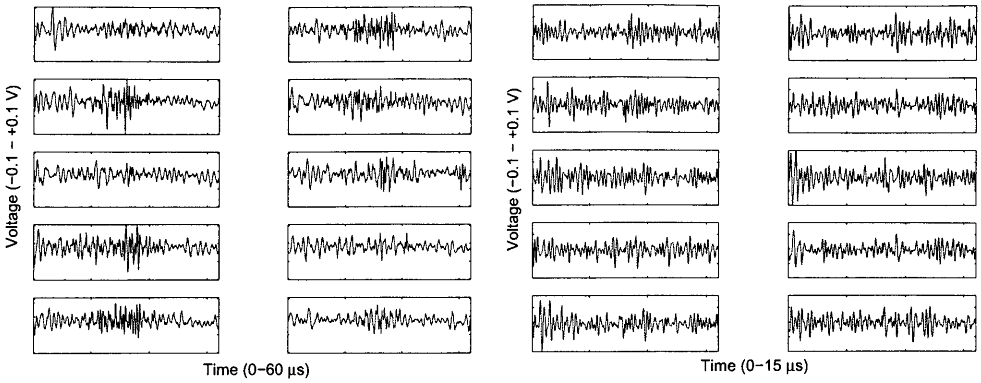
Figure 3-7: Windowed data from phantom experiments.

so the scattering is low when the pulse first encounters the tumor, then increases as the cross-section of the tumor grows wider, and then decreases once the wave moves beyond the center of the tumor.

As mentioned in Section 3.1, there were many experiments where the tumor grew in a location where it was not visible to our transducer. These experiments needed to be thrown out in order to develop a successful classification algorithm and so we kept the ten most visible tumors experiments. The last measurement of each of these experiments, when the tumor has reached a size of 0.3 cm^3 is shown in Figure 3-6.

The first issue was to align and window the data. When a doctor tries to determine if a suspicious mass is a tumor, he typically focuses attention on a particular mass and its surrounding tissue, ignoring the rest of the image. Sometimes, the properties of the surrounding tissue are important and should not be ignored. However, the algorithm attention should be somewhat limited to ease the detection task. The center for each window is set to be the maximum value of the wave. The window is then set to be wide enough to include the largest tumor size. The windowed data is shown in Figure 3-7.

The next step is to add the effect of tissue scattering to these waveforms. As mentioned is Section 3.1, unlike breast tissue, the scattering inside the gelatin is almost zero. Thus the experiments with no tumor growing produced a signal that was



(a) Scatter added to windowed data (b) Scattering alone

Figure 3-8: Windowed data with scatter added

never larger that ± 0.03 V. Detecting tumor growth in this environment is a trivial task. To make the task more difficult and more like the actual task of detecting a tumor in real breast tissue, we added the effect of scattering to the windowed waveforms. The scattering waveforms were generated using PZFlex, a finite element ultrasound simulation. The details of their generation will be discussed in Chapter 4. The new waveforms with the additional scattering are shown in Figure 3-8(a). The simulated scattering alone is shown in Figure 3-8(b).

3.2.2 The Classification Algorithm

Nearest Neighbor Learning

One simple way to classify the data is to do Nearest Neighbor Learning, as described in Chapter 4 of [21]. The basic principles of Nearest Neighbor (NN) Learning are illustrated in Figure 3-9. NN is a form of supervised learning for classification. One starts with a set of data points that are classified into n different classes. When a new data point of unknown type is received, the distance between that point and every other known point (D) is calculated. The new point is then determined to be of the same class as its closest neighbor. A modification to NN is known as kNN, which simply stands for k Nearest Neighbors. With kNN, instead of just classifying

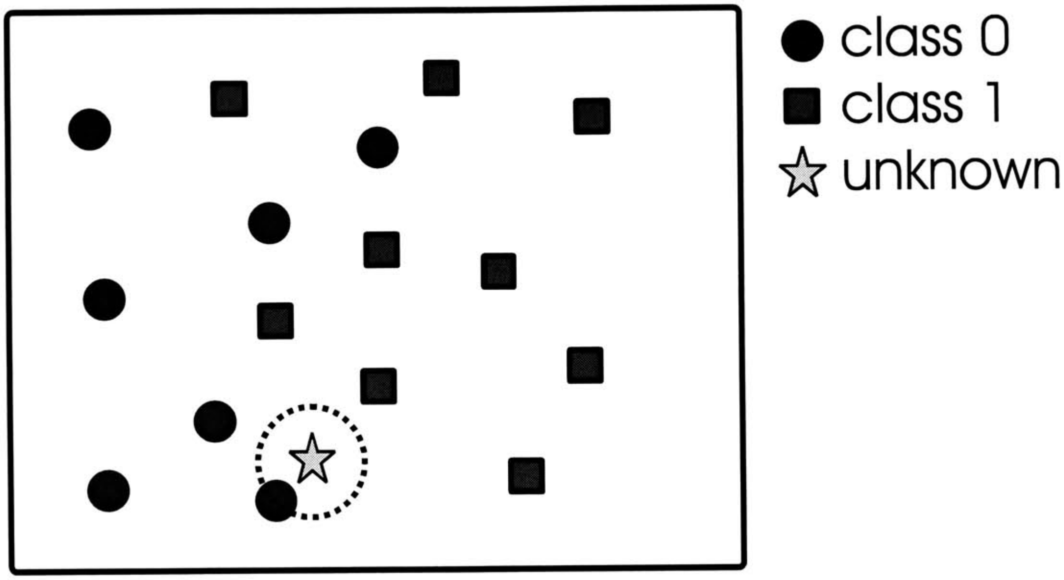


Figure 3-9: Illustration of Nearest Neighbor Learning. Here, since the unknown point is closest to a point in class 0, it is identified as being in class 0.

a point by its nearest neighbor, one looks at k neighbors, giving each neighbor one vote to determine the classification outcome. One can also choose to give all points equal votes, or to weight each vote by the distance of the unknown point from each of its neighbors.

The most obvious choice for a distance metric is the Euclidean distance, shown in Equation 3.1. But, there are others, including the L_k norm shown in Equation 3.2. A more in depth discussion can be found in [21]. [22] discusses an optimal distance measure for nearest neighbor classification.

$$D(a,b) = \left(\sum_{i=1}^{N} a_i^2 - b_i^2 \right)^{1/2} . \tag{3.1}$$

$$L(a,b)_k = \left(\sum_{i=1}^{N} a_i^k - b_i^k \right)^{1/k} . \tag{3.2}$$

When implementing this algorithm on the experimental data, one can represent each waveform as a vector ($\mathbf{v_i}$) in N dimensional space, where N is the number of points in the waveform. Recall that in each experiment, approximately 100 waveforms were stored as the tumor grew from nothing to its full size. For our detection algorithm, we chose to look only at 6 waveforms per trial, where the tumor volume increased by .06 cm^3 each time. To capture the entire growth history of the tumor in a single data point, one only has to concatenate a series of $\mathbf{v_i}$ vectors, such that if $\mathbf{v_i}$ is a row vector representing a single waveform, then $\mathbf{V} = [\mathbf{v_1}\ \mathbf{v_2} \ldots \mathbf{v_w}]$. The dimensionality of $\mathbf{V}$ is now at most $6N$. Each individual waveform is approximately 15 μs long and the sampling rate was 50 MS/s, so $N = 750$ and $6N = 4500$.

The first test of this algorithm was to look only at the final waveform, $\mathbf{v_n}$, the "image" of the tumor once it has reached full growth. To determine the class of a particular experiment ($\mathbf{V_i}$) we began with m experimental data points $\mathcal{V} = [\mathbf{V_1} \ldots \mathbf{V_m}]$, and eliminated $\mathbf{V_i}$ from the set to achieve $\mathcal{V}^i = [\mathbf{V_1} \ldots \mathbf{V_{i-1}}\ \mathbf{V_{i+1}} \ldots \mathbf{V_m}]$. We then used the Euclidean distance metric to find the point or points in $\mathcal{V}^i$ that were closest to $\mathbf{V_i}$ and identified $\mathbf{V_i}$ as being from that class. Using this method for evaluating a classification scheme is known as the Leave One Out method, as the training set includes all data points except for the one being tested.

The results of this test were very poor. The best classification scheme gave an accuracy of 54%. Interestingly, it occurred when $k = 1$. As k increased or as more waveforms were added to V, the accuracy dropped to 50%, which was no better than flipping a coin. This poor performance of this algorithm illustrates what is known as the curse of dimensionality, which states that as data takes on higher and higher dimensions, the number of possible classification boundaries increases very quickly. Right now, the problem we are trying to solve has from 750 to 4500 dimensions but we have only 20 data points with which to fill this space. One solution is to try and gather more data, but this is an infeasible solution as it would take at least thousands of points to fill this space. The task becomes even more difficult when we start using T transducers, which adds another multiplicative factor to the dimensionality of our problem. An alternative solution is to use data reduction techniques that reduce the number of dimensions in our problem.

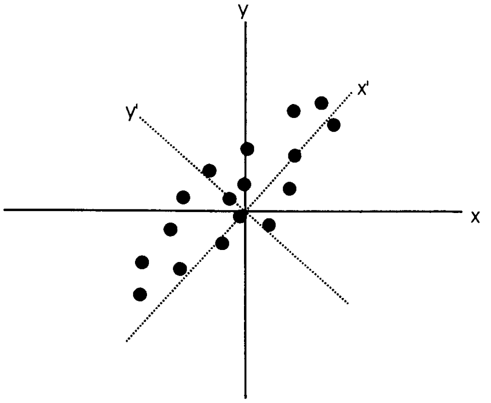


Figure 3-10: PCA Illustration.

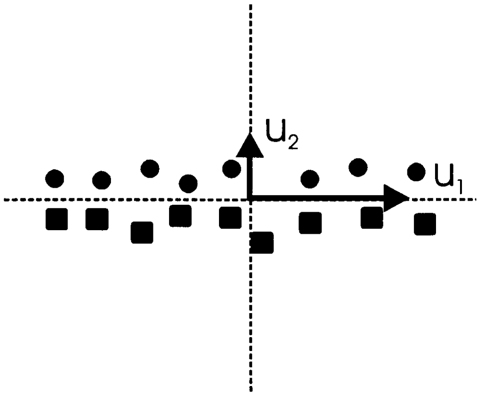


Figure 3-11: Sometimes, PCA is not good for classification. In this toy example, if PCA is used and u_2 is thrown out, classification will only be 50% accurate, whereas if u_1 is thrown out instead, classification will be 100%.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a technique that is commonly used for dimension reduction. An introduction is provided in [23]. It can be used for a wide variety of tasks including face recognition as described in [24]. The essential function of PCA is illustrated in Figure 3-10. The principal component of the data in Figure 3-10 is in the direction of maximum variance, x'. If we want to represent each point with only one dimension instead of two, PCA tells us the vector along which this dimension should lie in order to minimize the error between the reduced representation and the actual representation.

The principal components are calculated by taking the eigenvalues (λ_j) and eigenvectors (u_j) of the covariance matrix (Σ) of the data. When there are fewer data points (m) than there are dimensions, as is often the case, the covariance matrix is singular and has rank m. Thus there can be at most m principal components. The eigenvectors with the largest eigenvalues are the most significant principal components. To project v_i onto our new space, one simply calculates $v_i \cdot u_j$ for all j.

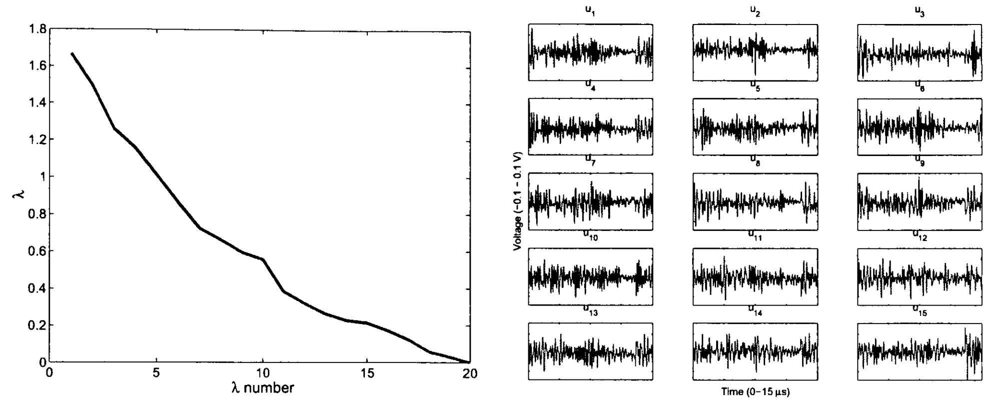
Some intuition for why PCA would improve the performance of our NN algorithm can be seen by revisiting Figure 3-6. We see that at the beginning and end of each waveform, there is just noise. The values of waveforms at these points are completely useless, but the NN algorithm uses them when computing the distance between data points. PCA can reduce the role that these points play in calculating the NN distance.

PCA is not the only data reduction technique. [25] explains that PCA is specifically designed to yield the minimum least squared error when representing data and it is not necessarily designed for classification, as illustrated in Figure 3-11. In the figure, we see that the least important dimension according to PCA is actually the most important dimension when it comes to classification. [25] discusses alternate data reduction techniques. For example, one scheme looks at the mean of the two different classes and the primary dimension points along the line connecting the means of the two different classes. The problem with this simple reduction scheme is that for two classes, all the data gets reduced to a single dimension which is often not enough for adequate data representation. Thus more sophisticated methods are also used. [25] then uses some different techniques to classify 21 different data sets. PCA averages 80.1% accuracy while the best alternative technique averages 82.4%. Thus even though PCA may not be an optimal form of data reduction for classification, it appears to still perform very well in general. Another option for data reduction would be to use wavelets, which are often used for decomposing time dependent waveforms

A plot of the principal components of $\mathbf{v_n}$ in order of biggest to smallest is shown in Figure 3-12(a). The orthonormal vectors that correspond to these components are shown in Figure 3-12(b). Notice how the principal component vectors have similar features to those in Figure 3-7. Both the actual data and the first principal components have peaks in the middle. The less important principal components look more like noise.

3.2.3 Results

kNN was performed with $k = [1,3,5,7]$ on the PCA'd data. The data entered into the PCA was of the format $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$ where i indicates the volume of the tumor. The results are shown in Figures 3-13 and 3-14 for $k = 1$ and $k = 7$ respectively.



(a) Principal Components

(b) Principal Component Vectors in order of λ magnitude

Figure 3-12: PCA of v_n

These plots show the central result for this part of the thesis. Notice that at 0 volume, the changes of detecting a tumor are around 50%, which is no better than flipping a coin. As the tumor grows and more data is acquired, the accuracy of the algorithm improves. The rate of improvement depends upon the number of principal components used to represent the data. The accuracy for $k = 1$ reaches 100% when the tumor reaches a full size of 0.3 cm^3 and p = 2-5. When $k = 7$, the maximum accuracy is 95%.

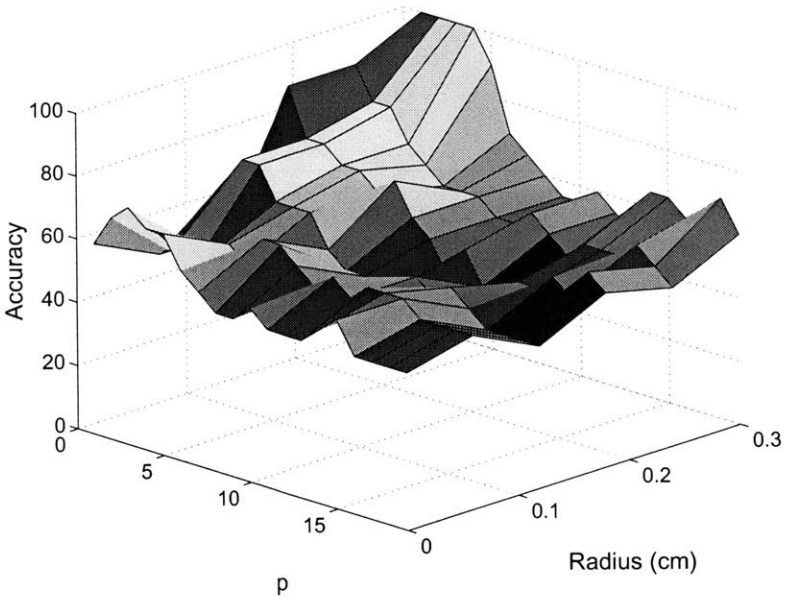


Figure 3-13: The results from running the detection algorithm on the processed experimental data where $k = 1$. p is the number of principal components used in the kNN algorithm. $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$ where i indicates the volume of the tumor in the plot.

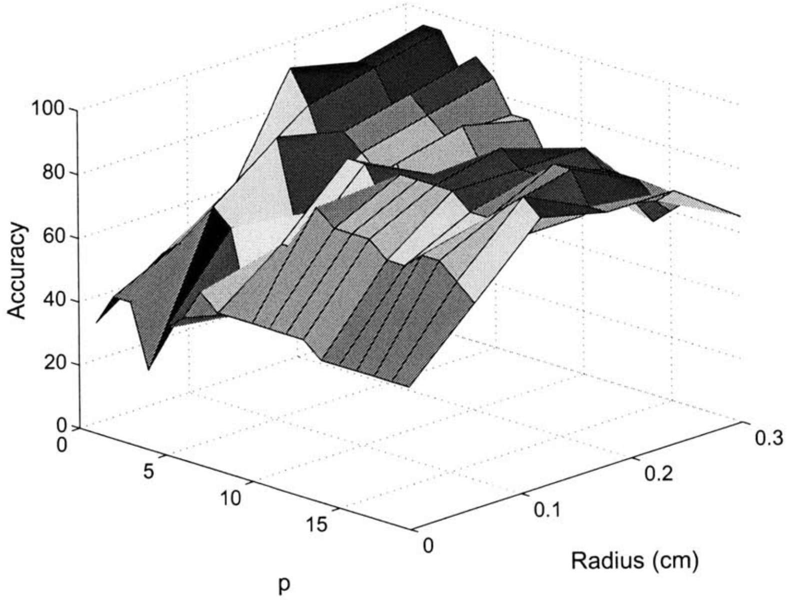


Figure 3-14: The results from running the detection algorithm on the processed experimental data where $k = 7$. p is the number of principal components used in the kNN algorithm. $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$ where i indicates the volume of the tumor in the plot.

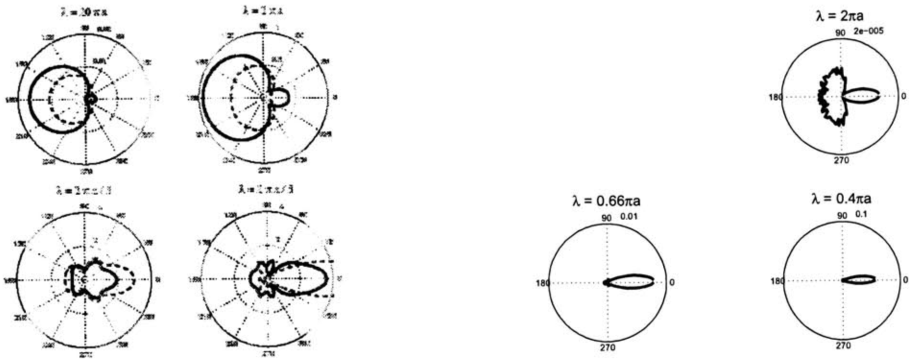
Chapter 4

The Simulation

4.1 The Model

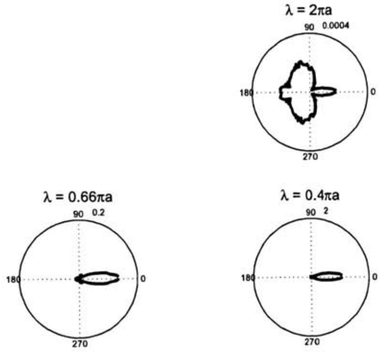
As mentioned in Section 3.2, a finite element simulation was used to generate the scattering data that was added to the experimental data in Figure 3-8(a). An additional benefit of developing a finite element model is that we have the advantage of completely controlling the geometry and we can run thousands of experiments with almost no effort. The package that we used was PZFlex, a linear, finite-element ultrasound simulation package made by Weidlinger Associates. To check the validity of the software, we simulated a plane wave scattering off a glass sphere. The results are shown in Figures 4-1 and 4-2. The following figures compare the analytical solution to data taken from the computer simulation. Notice how the qualitative shapes of the figures agree. Spheres that are smaller than a wavelength scatter most of the power back toward the source, while spheres that are larger than a wavelength scatter most of the power away from the source. Also notice that the total scattered power as a function of the sphere radius (a) in Figure 4-2 matches the theory reasonably well. The major discrepancy between the simulation and the theory is the roughness of the simulated curves when compared to the smooth analytical solution in Figure 4-1. We believe that this is due to the discrete nature of the simulation.

We chose to look only at 2D simulations because we believe they provide the same intuition as 3D simulations, but are much less computationally intensive. In



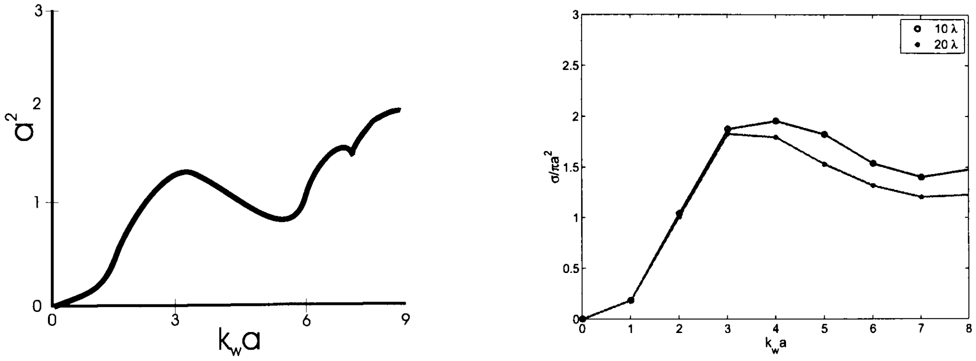
(a) Analytical Solution from http://dukemil.egr.duke.edu/Ultrasound/k-space/bme265.htm. The solid line is for an elastic sphere and the dashed line is for an inelastic sphere. The simulated spheres in the other figures are all for elastic.

(b) PZFlex simulated solution at 0.5 MHz with 25 n/λ



(c) PZFlex simulated solution at 2.0 MHz with 25 n/λ

Figure 4-1: σ_d for a plane wave scattering off a glass sphere in water. The source is located at 180°. Note that no simulations were run for $\lambda = 20\pi a$ because the mesh size required to run these simulations with a reasonable looking sphere would require a prohibitively long simulation time.



(a) Analytical Solution from http://dukemil.egr.duke.edu/Ultrasound/k-space/bme265.htm

(b) PZFlex simulated solution. The distance at which the wave was measured equal to 10λ and 20λ.

Figure 4-2: Scattered Power from a glass sphere in water. a is the sphere diameter, and k_w is the wave number of water

a typical simulation a tumor with a = 6 mm was be placed 2 cm from the source, and so a 3 × 1.5 cm window was used, as shown in Figure 4-3. For a 1 MHz signal with a wavelength of 1.5 mm, the window is 20 × 10λ. Also, typically 25 elements/λ were needed for reasonable results, so the geometry contains 500 × 250, or 125,000 elements. These simulation took 4 minutes on a Pentium III, 1 GHz computer. If we were to switch to 3D, we would need 500 × 250 × 250 elements, thus increasing the time required by a factor of 250 to 16.7 hours. Since we needed to run hundreds of simulations for our learning algorithms, 3D was clearly infeasible.

There are some important qualitative differences between 2D and 3D. Mainly, the intensity of a cylindrical wave decreases as $1/\sqrt{r}$, while a spherical wave decreases as $1/r$, where r is the distance from the scatterer. Thus the magnitude of returning waves in a 2D simulation will typically be larger than that of a 3D simulation. However, when we model the scattering of the environment, the "noise" from this environmental scattering is also be artificially increased, thereby negating the benefit from increased tumor scattering.

The geometry for the simulations is shown in Figure 4-3. The tissue indicated in the figure is either glandular tissue or a tumor. The boundaries of the simulation were

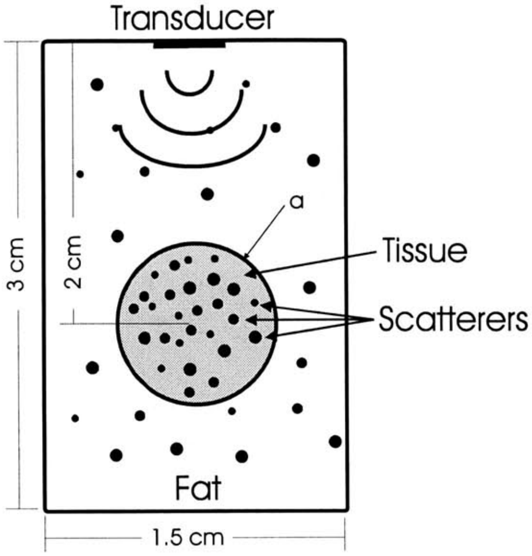


Figure 4-3: Simulation Geometry. a is the radius of the tissue

perfectly absorbing. The transducer was modelled by simply specifying a pressure value at a particular region of the model. The input pulse was the first derivative of a gaussian, $\frac{d}{dt}\frac{1}{2\pi\sigma}e^{-t^2/2\sigma^2}$, where σ is the spread. The input pulse is shown in Figure 4-4. σ was chosen so that the frequency would be centered at .75 MHz and less than 10% of its peak value at 2 MHz. Data was "acquired" by simply recording the pressure at the point of interest. This simulation completely neglects the transfer function of the transducer and any coupling effects between the transducer and the breast. These can be added later.

Parameters c and ρ were assigned according to Table 2.1. Attenuation was neglected. It increases the speed of the simulation by a factor of 2 by removing the need to calculate extra variables. At 1 MHz, attenuation through fat is .158 dB/cm. Thus even after a wave travels 4 cm, it has only attenuated .63 dB, resulting in a loss of only 8.4%.

The backscatter coefficient was a value that could not easily be programmed into the simulation because it comes from the the geometry of the structure itself. As discussed in Chapter 2, backscatter can occur on very small scales and is very difficult to model accurately. Our model for backscatter was inspired by Madsen's

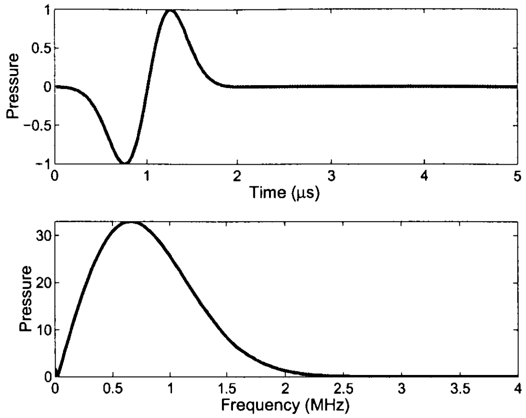


Figure 4-4: Input pulse from simulated transducer.

attempt to experimentally simulate breast tissue in [3], where he uses oil droplets and glass beads to cause acoustic scatter. Inside our tumor, we placed tiny spheres with a radii between 30 μm and 120 μm. We placed 20 spheres/mm^2 in glandular tissue, 2 spheres/mm^2 in the tumor and 1 sphere/mm^2 in the fat. This was to match the property that glandular tissue typically scatters 10 times more power than tumors and tumors scatter twice as much power as fat. The material properties of the scatterers were chosen to be $c = 850$ m/s and $\rho = 1350$ kg/m^3. We chose to use a c and ρ that is smaller than that of the fat because a higher c would have needed to be higher than the fastest c of the simulation. This means the shortest λ would be shorter and a finer mesh would have been required for the simulation, thus increasing the simulation time. The specific values were chosen by hitting an area of these scatterers with a plane wave and measuring the returning power. We can modify Equation 2.2 for 2D, as shown in Equation 4.1, where A is the cross-sectional area of the scatterer, l its length out of the plane, and $\mathcal{P}$ is force per unit length. Running the simulation with the above values for c and ρ and then measuring $\mathcal{P}$ allowed us ta calculate that $\sigma = 6.3 \times 10^{-5}$ cm^{-1}Sr^{-1}.

$$\sigma = \frac{1}{Al}\frac{\mathcal{P}(\theta,\phi)l}{I_0} = \frac{1}{A}\frac{\mathcal{P}(\theta,\phi)}{I_0} \tag{4.1}$$

Of course, this is probably a grossly inaccurate model of how scattering occurs in

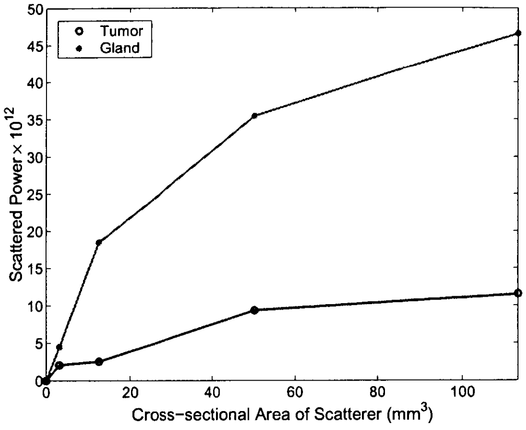


Figure 4-5: Scattering in the PZFlex simulation

biological tissue. But, it is a place to start. One artifact to be aware of is that, as mentioned earlier, we typically used 25 elements per wavelength, so at 1 MHz, the element size was 60 μm. The smallest scatterer also had a diameter of 60 μm, and so it takes up exactly one element in the simulation. The largest scatterer takes up 14 elements. These scatterers only vaguely resemble spheres. While biological tissue does not have these sharp boundaries, it is also not perfectly spherical. Again, more sophisticated model can be used later.

Figure 4-5 shows the scattering properties of our simulated tissue as a function of cross-sectional area. According to Equation 2.2, we would expect the scattered power to increase linearly with volume. Since we are using a 2D simulation, all geometric features have infinite length in the third dimension and volume is directly proportional to the area. What we see, however, is that the scattered power increases much slower than linearly. One possible reason for this trend is that the scatterers were not hit with a plane wave, but rather with a wave from a transducer of finite aperture. When the tissue of interest is significantly larger than the width of the beam, the effective volume hit by the pulse only increases with r instead of r^2, leading to a slower increase in the scattered power. Another feature to note is that while the scatterer density in the glandular tissue is 10 times greater than in the tumor, the power scattered by the glandular tissue is only a factor of 4 greater than the power scattered by the tumor . This discrepancy is due to the fact that the scatterers are added digitally and placed in random locations that do not depend on the position of the other scatterers. Thus at high scatterer concentrations, scatterers are likely to overlap.

As mentioned in Section 3.2, this simulated scattering was added to the experimental data before the detection task. The data was generated by simulating a wave propagating through fat and measuring the scattered power received. This wave was then compressed in time by a factor of 5 because the experiments were run using 5 MHz transducers while the simulation operated near 1 MHz. The scattered pressure wave was also scaled by an arbitrary amount before being added to the experimental data. If the scale factor was too high, it was impossible to see the vegetable shortening signal, but if it was too low, the detection task was too easy. Calculating the correct value for the scale factor requires knowledge of the transducer efficiency and the scattering properties of vegetable shortening, both of which were unavailable at the time of this publication.

4.2 Results

Data from the simulations are shown in Figure 4-6. This data, like the experimental data, was also aligned and windowed, but this time the data was aligned using the center of mass of the absolute pressure, rather than the maximum. This change in the windowing method was used because the scattering from these simulations had a relatively flat envelope compared to the experimental data. No additional scattering needed to be added because scattering was present in the model. This data was then entered into the PCA and kNN algorithm from 3.2. The PCA'd data was of the form $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$, where i indicates the current radius of the tumor. The results are shown in Figure 4-7 and 4-8 for $k = 1$ and $k = 7$ respectively . Using the simulation, we were able to simulate both tumor growth and glandular growth, and then try to distinguish a tumor from a gland (Figure 4-7(a) and 4-8(a)) and a tumor from an empty background (Figure 4-7(b) and 4-8(b)). Not surprisingly, it was easier to discern a tumor from an empty background of scatterers than to decide whether the growing tissue was a tumor or a gland. The plots show the same trends that were

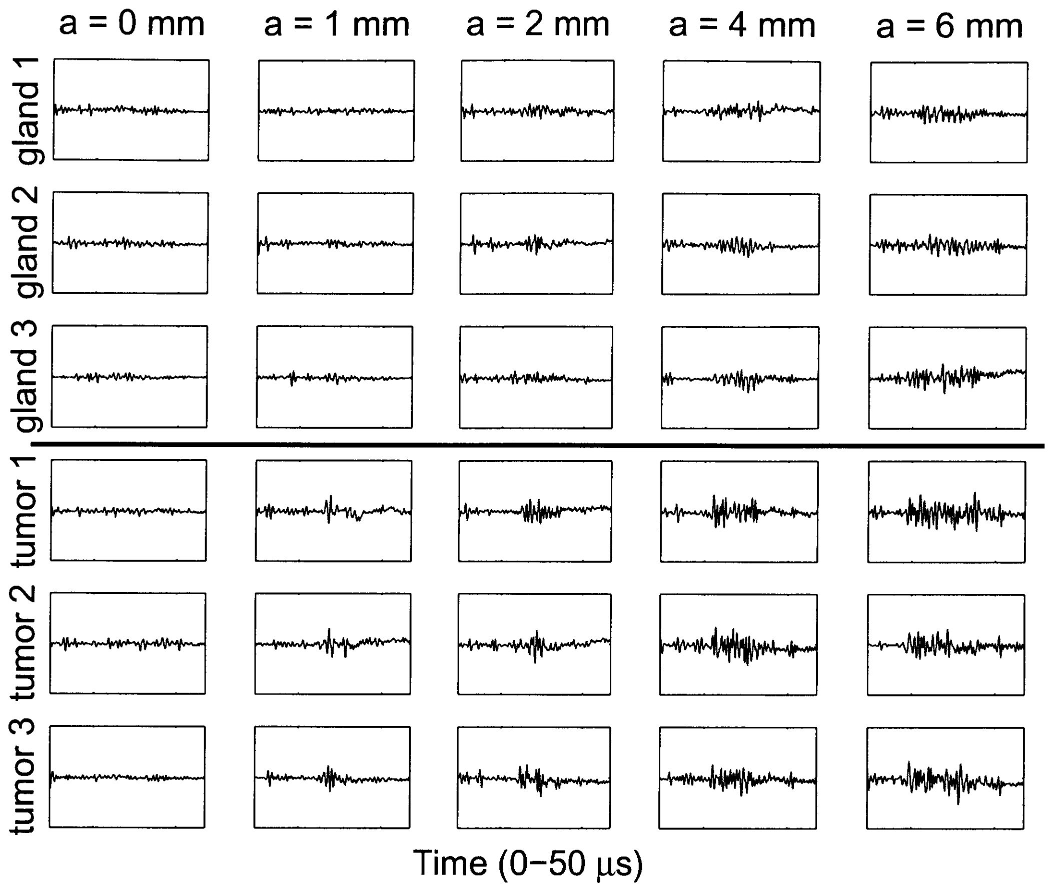
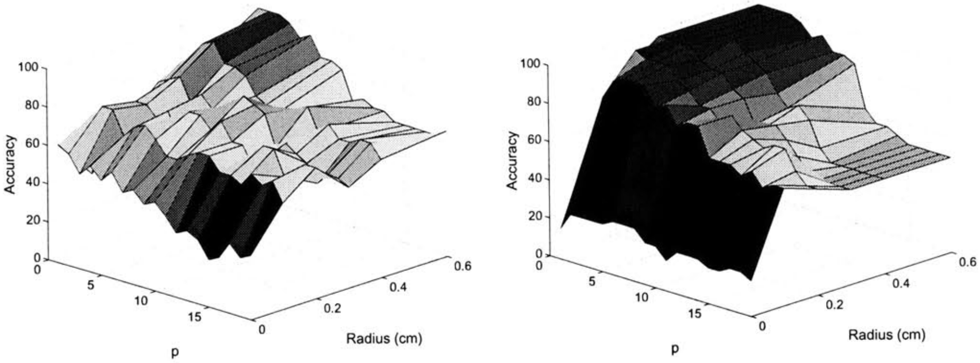


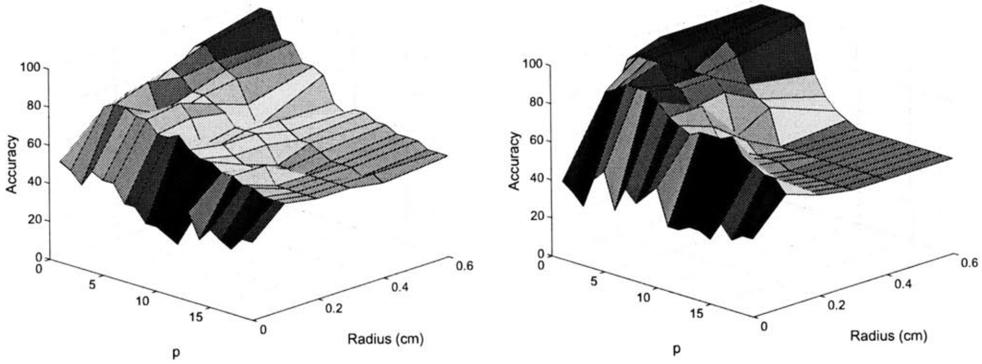
Figure 4-6: Pressure data from simulations. The columns are for different tissue radii. The rows are divided into two groups, one set of three for tumor simulations and a second set of three for glandular tissue. The random positioning of the scatterers is different for the waveforms within each group.

present in the experiment.



(a) Distinguishing tumor from gland (b) Distinguishing tumor from background

Figure 4-7: The results from running the detection algorithm on the simulated data. $k = 1$. p is the number of principal components used in the kNN algorithm. $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$, where i is the tumor radius shown in the plot.



(a) Distinguishing tumor from gland (b) Distinguishing tumor from background

Figure 4-8: The results from running the detection algorithm on the simulated data. $k = 7$. p is the number of principal components used in the kNN algorithm. $\mathbf{V} = [\mathbf{v_1} \ldots \mathbf{v_i}]$, where i is the tumor radius shown in the plot.

Chapter 5

Phase II Experiments

In an attempt to create an experimental environment in which we had more control over the acoustical properties of the materials involved, we devised a second experiment. This experiment ultimately failed to provide qualitatively reasonable results, but we have included its description for completeness.

These experiments made use of angioplasty balloons, growing inside a bath of water, as shown in Figure 5-1. The balloons were donated to us by Boston Scientific. The orientation and expansion profile of the balloon is shown in Figure 5-2. The angioplasty balloons were 1 cm in diameter and 6 cm long was filled with two different fluids: ethanol, and a 3 : 1 mixture by weight of water to hollow glass beads 2-20 μm in diameter. Ethanol was used because it had a different speed of sound (1207 m/s) and density (790 m/s), but very low backscatter. It simulated the tumor. The glass beads solution had a significantly higher scattering coefficient than the water, and thus it simulated the glands. Two sets of data were taken for d = [1,5] cm and Θ = [-45,-30,-15,0,15,30,45] degrees. Thus there were a total of 28 sets of data for each fluid. It should be noted that this positioning was not precise, and there was probably an estimated 20% error. The purpose of this round of experiments, however, was just to obtain a wide sampling of data and then try to separate the different classes fluid from each other. Thus it was not necessary to achieve precise, repeatable alignment.

The experiments consisted simply of filling the syringe with the desired fluid and then filling the balloon by manually depressing the plunger of the syringe. A mea-

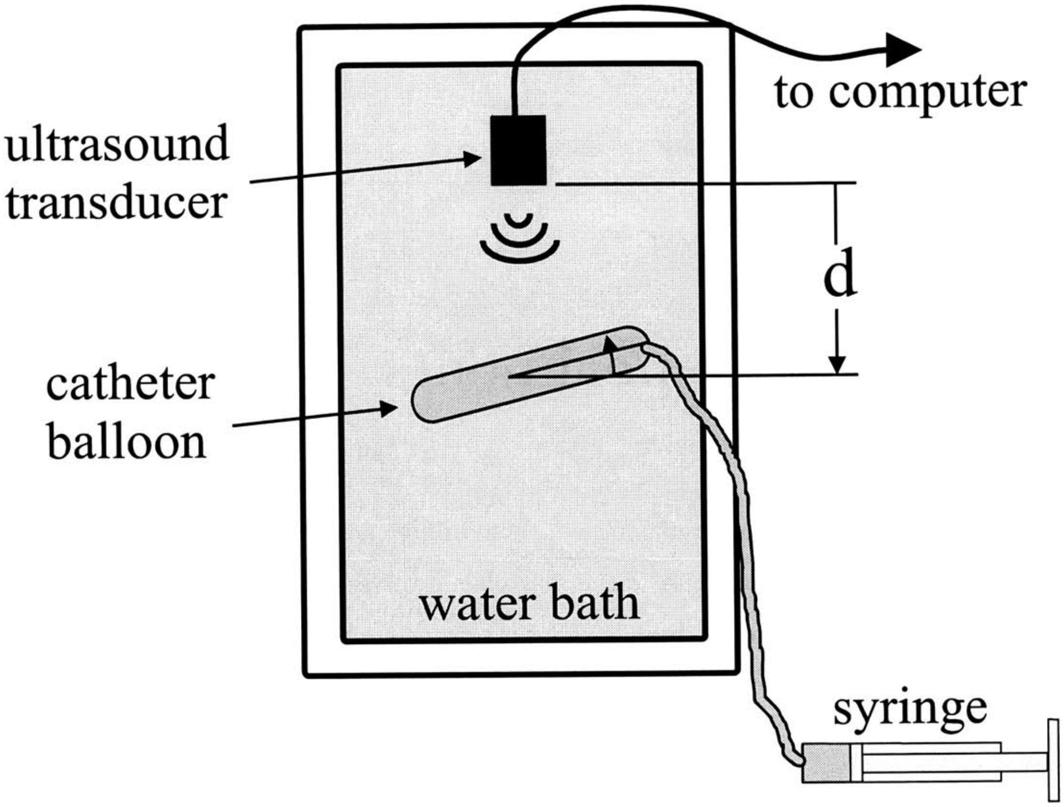


Figure 5-1: Top view of the phase II experimental setup. d is the distance from the transducer to the center of the balloon and Θ is angle between the balloon surface and the transducer surface.

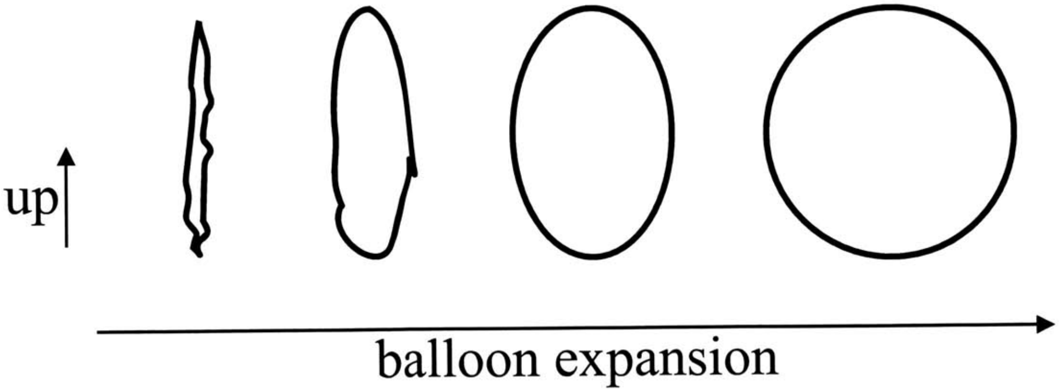


Figure 5-2: The cross section of the balloon as it expands.

surement was taken after every 0.3 cm^3 was injected until the balloon was full at 3.0 cm^3. Data from the Phase II experiments is shown in Figure 5-3. Notice how many of the plots have very sharp spikes. These spikes are due to reflections off the front and back interfaces of the balloon. It is these spikes that dominate the signal, and while the solution filling the balloon has some effect on the waveform, it was much too small.

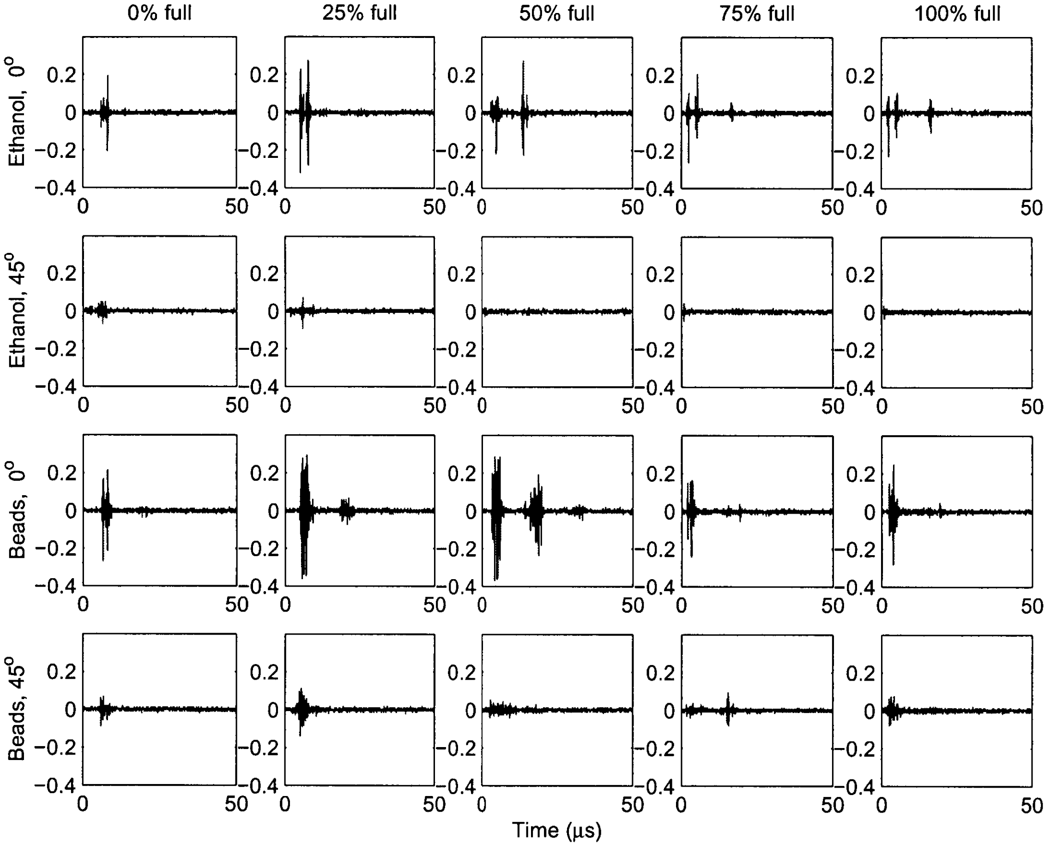


Figure 5-3: Example data from phase II experiments. Each row of data shows 5 waveforms for a single balloon as it inflates. The balloon was filled with either ethanol or the water/bead mixture and the angle of the balloon with respect to the transducer (Θ) was either 0° or 45°. This distance from the balloon to the transducer (d) was 1 cm.

The experimental data do not display the qualitative traits that we expected at the beginning. The question to be asked, then, is will these traits be relevant in actual, biological data and thus must be considered, or are these traits artifacts of our particular experiment that simply confuse the issue. We believe that the latter is true. We are not sure exactly what type of plastics the balloons are made of, but plastics in general have a $c \geq 2200$ m/s, while c for water and breast tissue ranges between $1430 - 1580$ m/s. When these experiments began, it was hoped that because the balloon walls were so thin, the scattering from the plastic would be negligible when compared to that of the bead-water solution. This may have been the case if these balloons expanded like a typical rubber balloon, such that their surface remained smooth throughout the experiment. However, as illustrated in Figure 5-2, the balloon surface was very wrinkled, causing significant scattering. Therefore, we decided to use an acoustic simulation package to generate data.

Chapter 6

Conclusions and Future Work

We have developed simple phantom experiments and a 2D model for simulating ultrasound propagation in breast tissue. In the experiments, we were able to detect the growth of a disk shaped tumor with 100% accuracy when it reached 0.3 cm^3, we achieved 100% accuracy when the tumor radius reached 0.6 cm.. The algorithm can also differentiate between tumor tissue and glandular tissue. The simulation and experiment should be compared only qualitatively because it is uncertain if the scattering model implemented in the experiments was scaled properly.

Of course with any simulation or phantom experiment, there is the possibility of ignoring important physics or real life details. For example, throughout this thesis we have assumed the ability to achieve precise transducer alignment. The the next step is to ensure correspondence between our simulation and some more realistic experimental results. We suggest working with Professor Madsen's group to design and fabricate phantoms for our study. Like the simulation, Madsen can control c, ρ, α, and σ of a material in order to simulate all types of breast tissue. Moreover, he can create multiple molds that are identical except for a change in the diameter of a single sphere, thus allowing for the simulation of tumor growth with a small shift in the surrounding tissue .

After establishing correspondence between the simulation and experimental data, the next step is to develop a more accurate model for how the breast changes from month to month. One way of viewing our problem is that we are trying to detect the tumor signature through environmental noise. By using randomly placed scatterers, we are in effect modelling this environmental noise as gaussian. If this noise is actually gaussian, the problem becomes relatively straight forward, and a matched filter is the best algorithm for extracting the signal from the background. However, if this noise is correlated, the task becomes more difficult and understanding how the breast changes with time and ignoring the benign changes becomes an important task. To our knowledge, no study of the monthly evolution of breast tissue has been undertaken to date. Thus, an additional task is to perform such a study where the evolution of the acoustic properties of tissue is monitored.

Once the acoustic evolution of the breast is understood and the simulation has been verified, many more simulations should be run. Our simulation detected the growth of a spherical tumor at a particular location with particular acoustic properties. A much more rigorous test, varying tumor shape, position, and material properties should performed to further test the detection algorithm. In addition, there are a wide variety of possible detection algorithms that can be tried, including Neural Networks, Support Vector Machines (SVM), and Hidden Markov Models (HMM). This thesis primarily focused on Nearest Neighbor learning because it was easy to implement, had a fast run time, and performed well. Moreover no significant performances were seen when the Neural Net or SVM were tried on the experimental data.

The HMM, briefly descibed in [21], is of particular interest because it has successfully been used for speech recognition, which shares many similarities with our problem. In both applications, one measures the properties of a waveform and how it evolves with time in order to decide which class it falls in. In speech recognition, the vocal cords modulate the voice, causing it to change with time. The task of the HMM is to interpret how these modulations occur over time and then decide which word has been spoken. In tumor detection, the tissue scatters the input pulse, creating a waveform that changes with time. So possibly, a similar HMM may be used to interpret the acoustic waveforms. In addition, it may be possible to use a higher level HMM to interpret the monthly changes in the waveforms that are caused by tissue growth.

In conclusion, we have found results that suggest that monthly ultrasound measurement using precisely aligned transducers will be able to detect smaller tumors than those currently detected in mammography. It is recommended that further work be done to better understand how the acoustic properties of breast tissue change from month to month and that more realistic experiments be run to confirm the results presented here.

Part II

The Tradeoff between Shift Invariance and Depth Resolution

Chapter 7

Shift Invariance vs. Depth Resolution Introduction

Known imaging systems capable of resolving object structure along the axial dimension (*e.g.* confocal microscopes, interferometers, binocular vision systems) usually exhibit limited lateral shift invariance; *i.e.*, they possess a location-dependent impulse response. For example, the confocal microscope [26], shown in Figure 7-1, achieves depth resolution via two means: (1) active illumination is focused on a specific object point, and (2) a pinhole placed at the detector plane rejects all out-of-focus light coming from the object. Both features severely limit the shift invariance of the system. If the pinhole is gradually opened up from its ideal zero diameter, the amount of lateral image information allowed through proportionally increases, whereas the depth-resolving capability decreases. In the limit of the pinhole becoming as wide as the field of view, axial imaging capability is essentially eliminated but that system has become perfectly shift invariant.

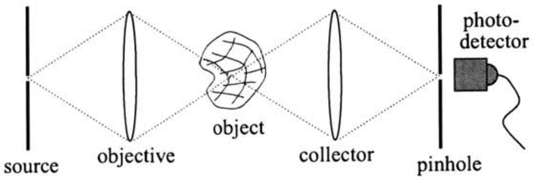


Figure 7-1: Confocal microscope.

The imaging process can generally be divided into two steps. First, projections of the object of interest are acquired. Because all detectors are limited to at most two spatial dimensions, the projections are also limited to 2D. To acquire a full 3D image, a second step is required, fusing multiple projections to form a single image. For example, a confocal microscope obtains projections that contain information about a single point within an object. By scanning along all three spatial dimensions, many point projections can be combined to form a full 3D representation of the object. Other examples of imaging systems that behave in such a way are radar, which integrates a series of measurements taken over time, and computed tomography, which uses the Radon transform to integrate projections of an object into the image or a slice from the object. Most good imaging systems will be shift invariant once the final image has been created. In this thesis, we are concerned with the behavior of individual projections.

We conjecture that lateral shift invariance and depth resolution are coupled in any object projection. Thus, by measuring the amount of shift invariance one should be able to estimate the depth resolution capability of an imaging system. Perhaps more interesting is the possibility of trading shift invariance off for depth resolution when designing an imaging instrument.

Among optical elements, volume holograms provide the capability of "tuning" the shift invariance at will. For example, this property has been studied in the context of holographic correlators for optical pattern recognition [27]. Recently, it was shown that a volume holographic matched filter can replace the pinhole of a confocal microscope to provide depth selectivity [28].

Here, we are interested in quantifying the shift invariance *vs.* depth resolution trade-off for general optical systems. For this purpose, we define metrics Δr and $1/\Delta z$ of shift invariance and axial resolution, respectively, such that they can be applied to general imaging systems. We then apply the definitions to a diffraction-limited confocal microscope and a binocular system, and show that shift invariance and depth resolution exhibit opposite trends as the pinhole radius of the system increases.

Chapter 8

Metrics for Shift Invariance and Depth Resolution:

The axial, longitudinal, or depth direction $\hat{\mathbf{z}}$ (we use all three terms interchangeably) with respect to an imaging system is the direction of an optical axis, if one is defined. If more than one axes can be identified in the system (e.g., in the case of multiple cameras), then we define the "effective axial" direction $\hat{\mathbf{z}}$ as the average of the axes of the system. The case of multiple optical axes will be further discussed when we examine the binocular system.

Let $h(x, y; x', y')$ denote the lateral intensity impulse response of an arbitrary imaging system at a fixed depth z_0. (x, y) are the coordinates of the source impulse at the input plane and (x', y') are the coordinates at the output plane. We define the shift variance domain Δr based on the function

$$S(x, y) = \frac{\iint_{-\infty}^{+\infty} [h(x, y; x', y') - h(0, 0; x' - x, y' - y)]^2 dx'dy'}{\iint_{-\infty}^{+\infty} |h(0, 0; x' - x, y' - y)|^2 dx'dy'}. \tag{8.1}$$

We will refer to $h(0, 0; x' - x, y' - y)$ as the shifted on-axis impulse response of the system. S gives a measure of the shift invariance of the system when the object is at (x, y). The denominator of (8.1) serves as a normalizing factor so that S is independent of the energy incident on the detector. A typical plot of S is shown in Figure 8-1. It is easy to see that S always equals 0 when $(x, y) = (0, 0)$ because the

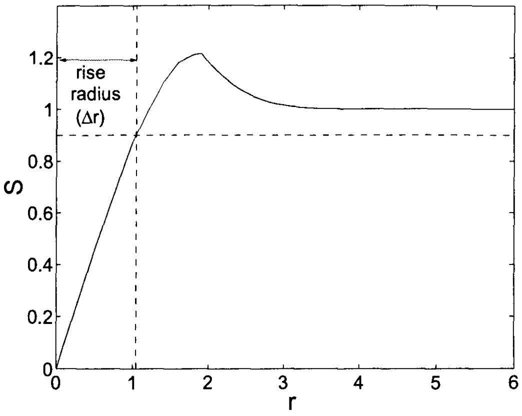


Figure 8-1: Shift Invariance for a Confocal Microscope. r is the radial coordinate normalized to the numerical aperture of the system, *i.e.* $r = \sqrt{v^2 + w^2}$

two terms in the numerator are the same. In a perfectly shift invariant system, such as a thin lens and detector with a fish-eye field of view, when an object shifts from $(0, 0)$ to (x, y), the image recorded by the detector also shifts by (x, y) and so the shifted impulse response exactly equals the image of the shifted object, making S=0 for all (x, y).

In a shift variant system, as the object moves away from the origin, the image measured at the detector can become distorted as in Figure 8-2. Thus the numerator in (8.1) diverges from zero. The more shift variant the system, the faster $S(x, y)$ diverges. The value of $S(x, y)$ may in fact increase beyond 1 if the overlap between the region where the impulse response and the shifted on-axis response becomes insignificant as shown in Figure 8-2. If we assume

$$\iint_{-\infty}^{+\infty} |h(x, y; x', y')|^2 dx'dy' \leq \iint_{-\infty}^{+\infty} |h(0, 0; x', y')|^2 dx'dy', \tag{8.2}$$

meaning that the intensity reaching the detector from a laterally shifted object is always less than or equal the intensity of the on axis object, then the impulse response and the shifted on-axis response may each contribute a total normalized intensity of at most 1. Therefore, $0 \leq S(x, y) \leq 2$ for a single detector. If there are multiple detectors, as in the binocular system, then $0 \leq S(x, y) \leq 2N$, where N is the number of detectors. In a system with a finite field of view,

$$\lim_{x,y\to\infty}\iint_{-\infty}^{+\infty}|h(x,y;x',y')|^2dx'dy'=0 \tag{8.3}$$

meaning the intensity measured at the detector approaches 0 as (x, y) approaches infinity, $S(x, y)$ converges to 1.

The metric we have created for shift invariance was designed to measure how fast S diverges from 0 as the object is moved away from the origin. The sooner the image becomes distorted, the more shift variant the system. The shift invariance metric we chose was the distance (Δr) from the origin required for S to reach 0.9, as shown in Figure 8-1. The value 0.9 was chosen because a typical shift variant system with a finite field of view will converge to 1, but for some systems, like the pinhole camera, S may not reach 1 until (x, y) reaches infinity. For a perfectly shift-invariant system, S never grows larger than 0 and so Δr is infinity. On the other hand, consider an infinitesimally small pinhole in the geometrical optics approximation. Such a system exhibits severe shift variance and its Δr approaches zero. For finite pinhole size and diffraction-limited imaging, Δr takes intermediate values, as we show later.

A simple way to measure axial imaging capability is to use the inverse of the uncertainty $1/\Delta z$ in determining the axial location of a point source around some reference depth z_0. This may depend on several factors, primarily detection noise but also image quantization due to a pixelated detector. Defined this way, $1/\Delta z$ also agrees with the common notion of "axial resolution." Both metrics Δr and $1/\Delta z$ generally depend on z_0; complete analysis of this phenomenon is beyond the scope of this paper, however. The $1/\Delta z$ for three different imaging systems will be discussed in more detail in the following section.

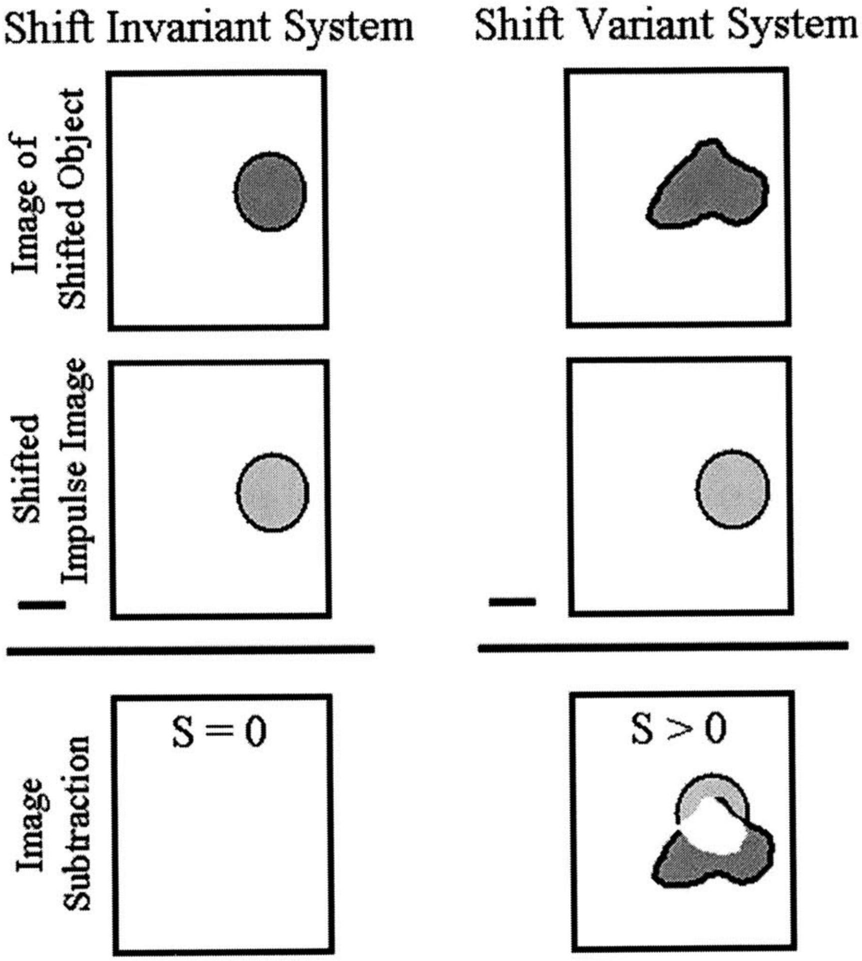


Figure 8-2: Shift Invariant vs. Shift Variant Systems. In the shift invariant system the image of a shifted object looks exactly like the shifted impulse response of that object. However, in the shift variant system, the image of a shifted object becomes distorted and no longer matches the shifted impulse response.

Chapter 9

The Behavior of S and $1/\Delta z$ in representative systems

9.1 The Confocal Microscope

Consider the confocal microscope system of Figure 7-1, equipped with a scanning mechanism, which allows it to acquire 3D data. This system is exceptional in that the shift variance and resolution do not depend on the depth z_0; this is because the system has a focal plane and the image data are acquired one point at a time. This property simplifies the understanding of the trade-offs that we want to discuss here.

The shift invariance data, shown in Figure 9-1, were generated from (8.1) using the in-focus lateral intensity impulse response of a lens, squared (to account for confocal illumination) and multiplied by the pinhole mask (radius a) as shown in normalized coordinates in (9.1).

$$h(v, v', w', w') = \left[\frac{2\,\mathrm{J}_1\left(\sqrt{(v-v')^2+(w-w')^2}\right)}{\sqrt{(v-v')^2+(w-w')^2}} \right]^4 \operatorname{circ}\left(\frac{r'}{a_n}\right). \qquad (9.1)$$

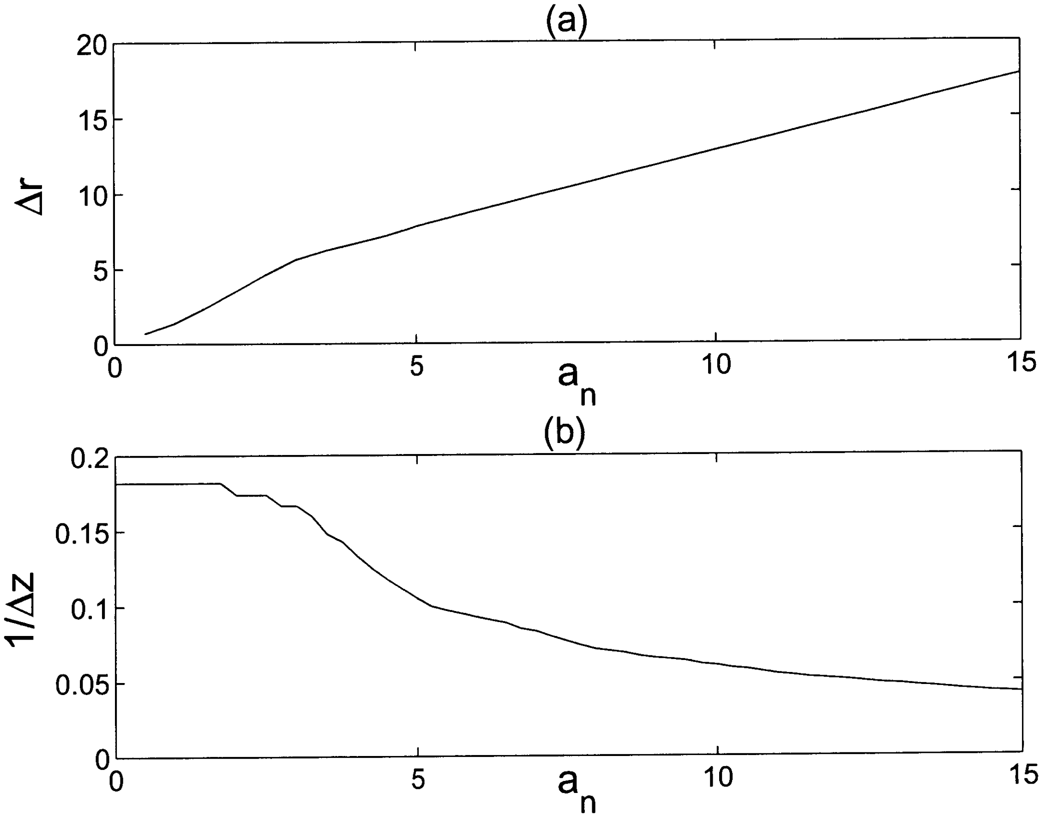


Figure 9-1: Confocal microscope tradeoff. (a) Shift Invariance vs. (b) Depth Resolution. Δr and Δz are normalized to the numerical aperture of the system

$$
\begin{aligned}
v &= \left(\frac{\pi}{\lambda(\mathrm{NA})}\right) x & \qquad v' &= \left(\frac{\pi}{\lambda(\mathrm{NA})}\right) x' \\
w &= \left(\frac{\pi}{\lambda(\mathrm{NA})}\right) y & \qquad w' &= \left(\frac{\pi}{\lambda(\mathrm{NA})}\right) y' \\
r &= \sqrt{v^2 + w^2} & \qquad r' &= \sqrt{(v')^2 + (w')^2} \\
 & & \qquad a_n &= \left(\frac{\pi}{\lambda(\mathrm{NA})}\right) a
\end{aligned}
$$

The depth resolution data of Figure 9-1 were generated using the integrated intensity derivation of [29, sec. 8.8.3]. The resolution Δz was defined as the full width at half maximum (FWHM) point. This is slightly pessimistic because it does not account for the improved signal-to-noise ratio at large pinhole diameters. With this caveat, the loss of shift invariance as depth selectivity improves is apparent from Figure 9-1.

9.2 The Binocular System

Now consider the binocular system of Figure 9-2. The origin for the object is selected to be location where the the focal planes of the two cameras intersect, and the optical axis for the system ($\hat{\mathbf{z}}$) is taken to be $(\widehat{\mathbf{z}_1} + \widehat{\mathbf{z}_2})/2$. The lateral axes are chosen such that $\hat{\mathbf{x}}$ lies in the same plane as $\widehat{\mathbf{z}_1}$ and $\widehat{\mathbf{z}_2}$ and is perpendicular to $\hat{\mathbf{z}}$. $\hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{y}}$.

This system is actually perfectly shift invariant in the $\hat{\mathbf{y}}$ direction. The shift invariance in the $\hat{\mathbf{x}}$ direction, however, has been destroyed. As the object moves in the $\hat{\mathbf{x}}$ direction, it moves out of the focal plane of each camera, thereby widening the point spread function (PSF) compared to that of the impulse response. In addition, as the object shifts on-axis, it moves closer to one camera and farther from the other. As the object moves closer to a camera, the magnification of the object by the lens

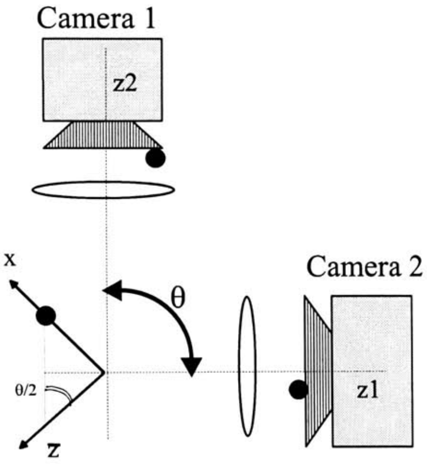


Figure 9-2: The Binocular System.

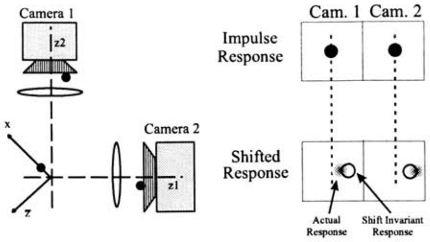


Figure 9-3: How the PSF shifts in the Binocular System.

also increases, thus increasing the displacement beyond what would be expected in a shift invariant system. The opposite is true when the object moves away from the camera. Both of these effects can be seen in Figure 9-3.

The corresponding shift invariance metrics for each detector, S_1 and S_2, are calculated individually, and S for the entire system is $S_1 + S_2$. In addition, since the system is symmetric, $S_1 = S_2$, making $S = 2S_1$. There are now two detectors in the system, so S has an upper bound of 4. This is twice the upper bound for a single detector. Because we ignore the finite field of view of the lenses in this example, S converges to this upper bound. A plot of S for a line segment that lies in the $\hat{\mathbf{x}}\hat{\mathbf{z}}$-plane and is perpendicular to the $\hat{\mathbf{z}}$ axis is shown in Figure 9-4. We ignore the $\hat{\mathbf{y}}$ dimension because the system is shift invariant in this direction. A plot of Δr vs. θ (the angle between the two cameras) is show in Figure 9-5. The derivation of the equations used to generate this figure is given in the Appendix.

There are two possible metrics for calculating the depth resolution. One takes into account the pixel size of the detector. With this metric, the depth is calculated by triangulation. Using simple geometry, we can see from Figure 9-2 that $\tan(\theta/2) = \Delta x/\Delta z$ and $\Delta x = M\Delta x'$, where M is the magnification of the lens. The resolution limit is when the pixel size (p) is equal to $\Delta x'$. Hence

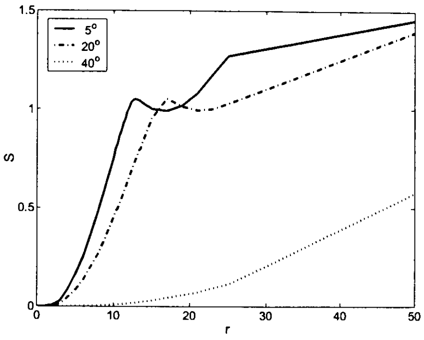


Figure 9-4: Shift Invariance for a Binocular System.

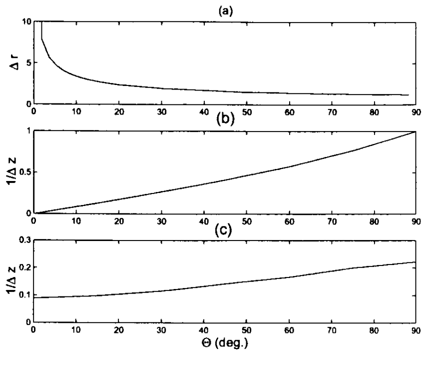


Figure 9-5: Binocular system tradeoff. (a)Shift Invariance vs. (b) Depth Resolution for triangulation and (c) Depth Resolution for PSF multiplication.

$$1/\Delta z = M \tan(\theta/2)/p. \tag{9.2}$$

The other possible metric for calculating $1/\Delta z$ is to compute the FWHM in the $\hat{\mathbf{z}}$ direction for the 3D PSF of the system. The PSF is generated by taking the PSF of one camera and point multiplying it by the PSF of the second camera. Figure 9-6 shows this point multiplication graphically. This second depth resolution metric is interesting because it does not depend on pixel size and is limited only by noise. Plots of both $1/\Delta z$ metrics are shown in Figure 9-5. Notice that when $\theta = 0$, the optical axes of the two cameras are parallel and the system performs no differently than if there was a single camera. It is perfectly shift invariant and has no depth resolution. As θ increases, the system becomes less shift invariant, while the depth resolution improves.

With the addition of more cameras, we point multiply by an additional camera PSF to get the system PSF and so depth resolution improves even more. This agrees with the intuition resolution should improve as the number of detectors increases. The addition of extra cameras also increases S, which is now $\sum_i S_i$, causing Δr to decrease.

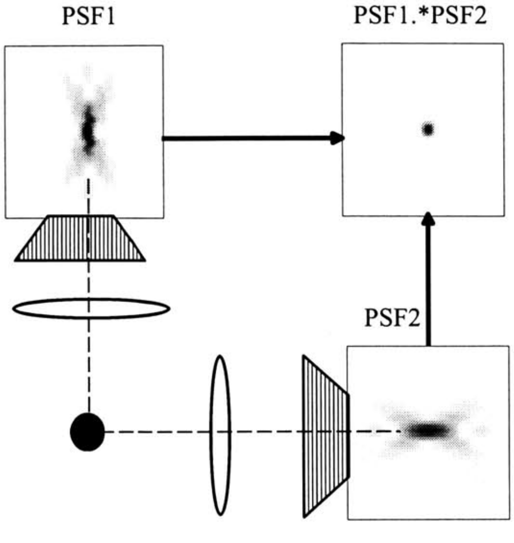


Figure 9-6: Point Spread Function Multiplication.

9.3 The Shear Interferometer

Now consider the shear interferometer of Figure 9-7 where an ideal point source creates a sinusoidal pattern on the detector. The intensity seen at the detector is given by

$$I(x,z) = 1 + \cos\left[\pi b\left(\frac{2(x-x_0)-h}{\lambda z}\right)\right], \tag{9.3}$$

where b is the shear introduced by a Bates interferometer or using a birefringent slab crystal [30]. From (9.3), we see that $\alpha \propto 1/z$ where α is the spatial frequency of the intensity at the output plane. Thus, as the point source moves away from the detector, the spatial frequency of the pattern on the detector decreases. Since the depth can be determined directly from α by using (9.3), the depth resolution can be calculated by just differentiating $I(x,z)$, which yields $\Delta\alpha/\Delta z \propto 1/z^2$ and by rearranging the equation, we get $1/\Delta z \propto 1/(\Delta\alpha z^2)$. According to the Nyquist criterion, $1/\Delta z \propto pz^2$, the minimum pixel size $p = 2/\Delta\alpha$.

Also from (9.3), we see that the phase of the pattern varies linearly with x and so the system is perfectly shift invariant. When the point source is close to the birefringent material, we have a shift invariant system with very high depth resolution and at first there appears to be no trade off at all. In this case, however, we have

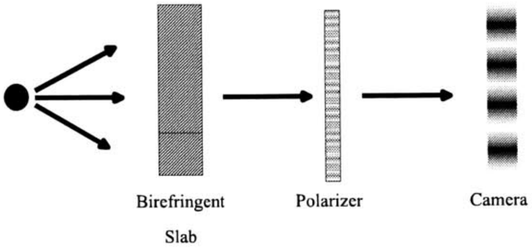


Figure 9-7: The Shear Interferometer

a new tradeoff between lateral ambiguity and depth resolution. When an object is close to the detector and f is high, there are a large number of positions in the lateral direction at which the object could be located, but when the object is far from the detector and f is low, there are only a few. This tradeoff is shown in Figure 9-8.

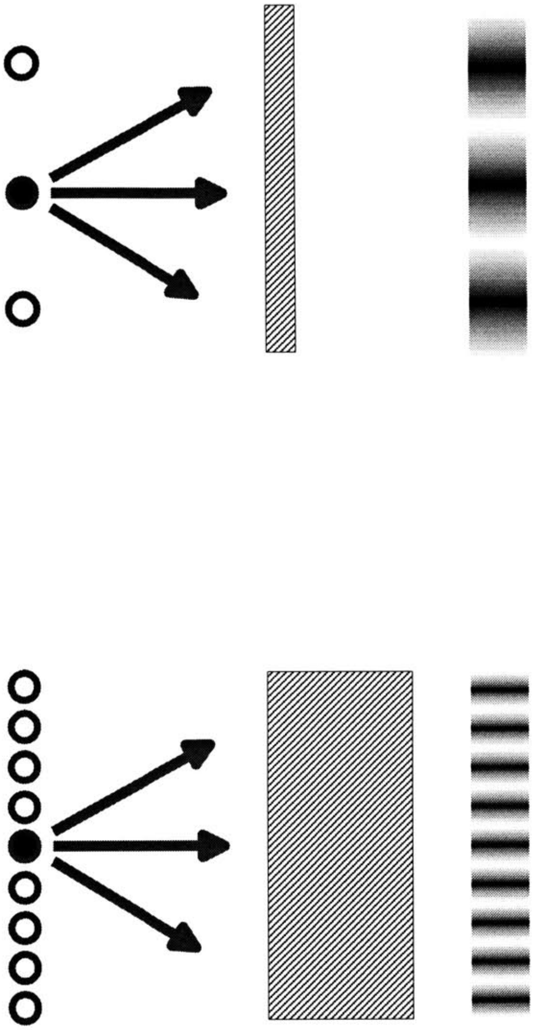


Figure 9-8: Ambiguity of the shear interferometer. In both pictures, the solid disk represents the actual position of the point source and the unfilled disks show other possible object locations that would have yielded identical patterns. A thin slab gives a low frequency interference pattern with poor depth resolution but has low lateral ambiguity, while a thick slab gives a high frequency interference pattern but has high lateral ambiguity

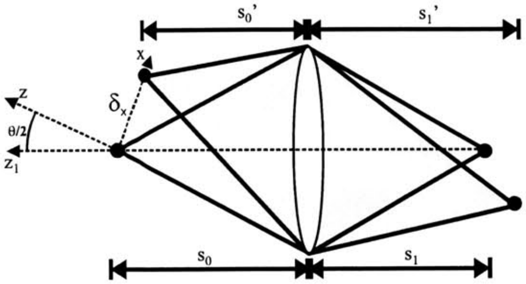
Chapter 10

Discussion and Conclusions

The competition among shift invariance and depth resolution is clear from the data of Figures 9-1 and 9-5 . It is also evident in many imaging systems that evolved in Nature. For example, humans can tolerate limited depth selectivity in favor of relatively large domain of shift invariance. This is because humans' cognitive capabilities can compensate depth and shape perception from other cues, such as object size, shading, texture and general knowlegde. We believe that similar exchanges can be applied to the design of artificial "smart" imaging systems which trade aspects of their image quality to maximize overall performance in cognitive tasks.

The fact that shift invariance and depth resolution are related suggests that there is an underlying property that many of these imaging systems share. Consider two Hopkins matrices. H_1 is simply the identity matrix, I. H_2 is H_1, but with its rows randomly permuted, such that all the rows are still orthogonal linearly independent, but the ones don't all fall along the diagonal. H_1 and H_2 both have the same mutual information, but H_1 is perfectly shift invariant while H_2 is highly shift variant. It seems, then, that one could prove that in many imaging systems, a Hopkin's matrix like H_2 cannot exist.

Appendix



We start with the Gaussian lens formula below, where f is the focal length and s_0 and s_1, are shown above.

$$\begin{aligned} M &= -\frac{s_0}{s_1} \qquad (1) \\ \frac{1}{s_0}+\frac{1}{s_1} &= \frac{1}{f} \qquad (2) \end{aligned}$$

If the object is shifted by a small amount, δ_x in the $\hat{\mathbf{x}}$ direction, the image shifts by $+\delta_x \sin(\theta/2)$, along the $\hat{\mathbf{z_1}}$ axis and $-\delta_x \sin(\theta/2)$ along the $\hat{\mathbf{z_2}}$ axis. For simplicity, we will define $\delta_{z_1} \equiv \delta_x \sin(\theta/2)$. Modifying the Gaussian lens formula yields

$$\begin{aligned} \frac{1}{s_0'}+\frac{1}{s_1'} &= \frac{1}{f} \qquad (3) \\ s_0' &= s_0 - \delta_{z_1} \qquad (4) \\ s_1' &= s_1 + \epsilon \qquad (5) \\ M' &= -\frac{s_0'}{s_1'} \qquad (6) \end{aligned}$$

Using a Taylor series expansion and assuming $s_0, s_1 \gg \delta_{z_1}, \epsilon$, we have

$$\frac{1}{s_0}-\frac{\delta_{z_1}}{s_0^2}+\frac{1}{s_1}+\frac{\epsilon}{s_1^2}=\frac{1}{f}. \qquad (7)$$

By subtracting (2) from (7),

$$\epsilon = \frac{s_0^2}{s_1^2}\delta_{z_1}. \tag{8}$$

Solving for M' in terms of δ_x, s_0, and s_1, gives

$$M' = -\frac{s_0 - \delta_{z_1}}{s_1 + M^2\delta_{z_1}} \tag{9}$$

Using a Taylor series expansion and again assuming $s_0, s_1 \gg \delta_x, \epsilon$, we can simplify M' to

$$M' = -\frac{s_0}{s_1}\left[1 - \frac{\delta_{z_1}}{s_0} - \frac{(s_0^2/s_1^2)\delta_{z_1}}{s_1}\right], \tag{10}$$

To simplify further, we introduce ΔM and A.

$$A = -\left[\frac{1}{s_0} + \frac{M^2}{s_1}\right] \tag{11}$$

$$\Delta M = M' - M = A\delta_{z_1} \tag{12}$$

Now, if the system were perfectly shift invariant, a shift of δ_x in the object plane would result in a shift of $M\delta_x$ in the image, but in reality we see a shift of $M'\delta_{z_1}$ which equals $M\delta_{z_1} + A(\delta_{z_1})^2$. Another way to express this property would be to say that in a shift invariant system, if an object moved at a constant speed in the $\hat{\mathbf{x}}$ direction, it would appear to move at a constant speed in the observation plane, but in a shift variant system, the object appears to accelerate. Moreover, the larger the angle between the cameras, the faster the acceleration will occur and the more shift variant the system. It should be noted that all these calculations for M' have been done for the camera that object moves toward when shifting in the $\hat{\mathbf{x}}$ direction. For the other camera, $M' = M - \Delta M$, and the object appears to decelerate rather than accelerate.

To calculate an approximate value for S, we start with a line of width W that is normal to $\hat{\mathbf{z}}$ and lies in the $\hat{\mathbf{x}}\hat{\mathbf{z}}$ plane. Since the system is completely shift invariant in the $\hat{\mathbf{y}}$ direction, we have chosen to reduce $S(x, y, x', y')$ to $S(x, x')$ for simplicity. Using geometrical optics, we estimate that this object forms a line of length $M'W\cos(\theta/2)$ on each camera. The response for this object is

$$h(x, x') = \text{rect}[M'(x)W\cos(\theta/2) - \Delta M(x)]. \tag{13}$$

A more precise calculation would involve using a point source and accounting for diffraction. S and Δr for this more complex system would exhibit the same qualitative behavior as our simple example.

Bibliography

[1] F. T. D'Astous and F. S. Foster, Ultrasound in Med. and Biol. **12**, 795 (1986).

[2] J. C. Bamber, in *Encyclopedia of Acoustics*, edited by M. J. Crocker (John Wiley and Sons, New York, 1997), pp. 1703–1726.

[3] E. L. Madsen, J. A. Zagzebski, and G. R. Frank, Ultrasound in Med. and Biol. **8**, 381 (1982).

[4] SEER Cancer Statistics Review (1973-1997).

[5] M. Moskowitz, in *Cancer of the breast*, 4 ed., edited by W. L. Donegan and J. S. Spratt (W. B. Saunders Company, Philadelphia, 1995).

[6] R. Sivaramakrishna and R. Gordon, IEEE Canadian Conference on Electrical and Computer Enginering **2**, 673 (1998).

[7] L. E. Kinsler and A. R. Frey, *Fundamentals of Acoustics* (Wiley, New York, 1950).

[8] K. K. Shung, in *Ultrasonic scattering in biological tissue*, edited by K. K. Shung and G. A. Thieme (CRC Press, Boca Raton, 1993).

[9] M. F. Insana and D. G. Brown, in *Ultrasonic scattering in biological tissuet*, edited by K. K. Shung and G. A. Thieme (CRC Press, Boca Raton, 1993).

[10] J. J. Faran, Journal of the Acoustical Society of America **23**, 405 (1951).

[11] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, New Jersey, 1968), pp. 418–439.

[12] M. E. Anderson, M. S. C. Soo, and G. E. Trahey, Ultrasound in Med. and Biol. **27**, 75 (2001).

[13] F. S. Foster, M. Strban, and G. Austin, Ultrasonic imaging **6**, 243 (1984).

[14] L. Landini, R. Sarnelli, M. Salvador, and F. Squartini, Ultrasound in Med. and Biol. **13**, 77 (1987).

[15] E. L. Madsen *et al.*, Ultrasonic imaging .

[16] F. A. Duck, Ultrasound in Med. and Biol. **28**, 1 (2002).

[17] F. W. Kremkau, *Diagnostic Ultrasound, Principles and Instrumentation*, 5 ed. (W. B. Saunders Company, Philadelphia, 1998).

[18] E. L. Madsen, J. A. Zagzebski, R. A. Banjavie, and R. E. Jutila, Med. Phys. **5**, 391 (1978).

[19] M. M. Burlew *et al.*, Radiology **134**, 517 (1980).

[20] E. L. Madsen, J. A. Zagzebski, and G. R. Frank, Ultrasound in Med. and Biol. **8**, 277 (1982).

[21] R. O. Duda, P. H. Hart, and D. G. Stork, *Pattern Classification*, 2 ed. (John Wiley and Sons, New York, 2001).

[22] R. D. Short and K. Fukunaga, IEEE Transactions on Information Theory 622 (1981).

[23] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, Oxford, 1995).

[24] M. Turk and A. Pentland, Journal of Cognitive Neuroscience **3**, 71 (1991).

[25] A. Tsymbal *et al.*, Proceedings of the 15th International FLAIRS Conference **IT-27**, 622 (2002).

[26] M. Minsky, Microscopy Apparatus, with United States Patent 3,013,467, granted 1961; T. Wilson, editor, *Confocal microscopy*, Academic Press, 1990; J. K. Stevens, L. R. Mills, and J. E. Trogadis, editors, *Three-dimensional confocal microscopy: volume investigation of biological systems*, Academic Press, 1994.

[27] M. Levene, G. J. Steckman, and D. Psaltis, Appl. Opt. **38**, 394 (1999).

[28] G. Barbastathis, M. Balberg, and D. J. Brady, Opt. Lett. **24**, 811 (1999).

[29] M. Born and E. Wolf, *Principles of Optics*, 7 ed. (Pergamon Press, London, 1998).

[30] W. J. Bates, Proc. Phys. Soc. **59**, (1947).