Title
Ultrasonic 3D Rangefinder on a Chip

Author
Przybyla, Richard J.

Publication Date
2013

Ultrasonic 3D Rangefinder on a Chip

by

Richard James Przybyla

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Engineering: Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernhard E. Boser, Chair
Professor Kristofer S. J. Pister
Professor Liwei Lin, Mechanical Engineering

Fall 2013

Ultrasonic 3D Rangefinder on a Chip

Abstract

Ultrasonic 3D Rangefinder on a Chip

by

Richard James Przybyla

Doctor of Philosophy in Engineering: Electrical Engineering and Computer Science

University of California, Berkeley

Professor Bernhard E. Boser, Chair

Optical 3D imagers for gesture recognition, such as Microsoft Kinect, suffer from large size and high power consumption. Their performance depends on ambient illumination and they generally cannot operate in sunlight. These factors have prevented widespread adoption of gesture interfaces in energy- and volume-limited environments such as tablets and smartphones. Gesture recognition using sound is an attractive candidate to overcome these difficulties because of the potential for chip-scale solution size, low power consumption, and ambient light insensitivity.

Our research focuses on building a 3D ultrasonic rangefinder system using batch-fabricated micromachined aluminum nitride (AlN) ultrasonic transducer arrays and custom CMOS electronics. The system uses pulse–echo time–of–flight to localize targets from their echoes. We use millimeter–wave ultrasound, which enables compact ultrasonic arrays which can measure range and direction to a target. The relatively slow speed of sound allows the use of low–speed, low–power readout electronics.

In this dissertation, we will present the design methodology for a prototype ultrasonic rangefinder system. We will show how the choice of basic system specifications affects the mechanical transducer design and the interface circuit design. We will present a physics-based model of an ultrasound transducer which accurately predicts device operation. We will present measured results from an ultrasonic 3D gesture recognition system which uses an array of AlN MEMS transducers and custom readout electronics to localize targets over a $\pm 45^o$ field of view up to 1m away. The 0.18μm CMOS readout ASIC comprises 10 independent channels with separate high voltage transmitters, readout amplifiers, and ADCs. Power dissipation is 400μW at 30fps, and scales to 5μW/ch at 10fps.

To my Grandparents: Kathryn, Phil, Bess, and Dick.

Contents

List of Figures

List of Tables

Acknowledgments

The rich variety of intellect, culture, and research that I've experienced at Berkeley has opened my mind to the possibilities that the future holds for us and the important role that technologists have in shaping that future.

While here, I have had the incredible fortune to meet many people far smarter and more driven than I, hailing from all corners of the planet and many different walks of life. Just like the Oscars, I'm sure I will miss many notable people who've helped me or enriched my experience here, and for that I apologize.

First and foremost I'd like to thank my parents Jim and Jill for their support and unwavering love. My parents gave me the best upbringing a kid could ask for and are the best friends an adult could have. My sister Jaime and my brother Chris can always be counted on for a good time when we get to spend time together and I've sorely missed them over the last five years.

I'd like to thank Bernhard for being a great advisor. I know that my work is ready for the public eye if it can get past Bernhard's legendary red ink, and although I may have grumbled about it at the time I've learned how valuable his feedback is to my research. I've learned to appreciate his style, efficiency, and technical depth and I hope I can emulate these features in the future, whether or not I get the chance to rejoin the academic world.

Bernhard's group has been a great place to come of age as an engineer. The wide–ranging lunch discussions that Igor Izyumin, Mitchell Kline, Dr. Octavian Florescu, and I have had over the years have been one of my favorite parts of going to work (especially some days...) and I hope they will continue for many years to come. Igor and Mitchell have always been ready to lend an ear to a problem and I hope we get the chance to work together in the future. I enjoyed learning the ways of Berkeley from Dr. Karl Skucha. Hao–Yen Tang saved me many times over on our tapeout and I am very thankful for his impressive skills. I've enjoyed getting to know Pramod Murali, Oleg Izyumin, Behnam Behroozpour, Dr. Paul Liu, Reut Krigel, Burak Eminoglu, Yu–Ching Yeh, Dr. Mischa Megens, and Dr. Mekhail Anwar and feel privileged to work in such a smart crowd. I've appreciated the advice received from Prof. Simone Gambini over the years and look forward to visiting him and Dr. Rikky Muller Down Under one day.

I'd like to thank Prof. David Horsley for being a great unofficial second advisor, and thank the researchers in his group including Stefon Shelton, André Guedes, Ofer Rozen, and Yi–Peng Lu for the great work they've done developing the transducers we use. I've enjoyed the many technical discussions that have come out of this great collaboration.

I'd like to thank Prof. Kris Pister for livening up the SWARM lab and for his helpful advice, Prof. Dick White for his work testing our systems with gas flow meters, and Prof. Elad Alon and Prof. Ming Wu for the teaching knowledge I've gleaned while teaching with them. I'd like to thank the staff in the department who do a great job fighting through the bureaucracy, especially the BSAC staff including Kim Ly and Richard Lossing. Mila MacBain gave great life to the SWARM lab while she was working here and I'm sure she will be missed for years to come.

I'd like to thank all the friends, roommates, and acquaintances I've had the pleasure to know while I was here. My flatmates Dr. Sobhan Seyfaddini, Aurelien Vidal, Dr. Diogo I–Have–Too–Many–Names Silva, Dr. Giulia Gubitosi, Ning Zhang, Steven Callender, Nitesh Mor, Amit Gupta, Amik Singh, and Pauline Loh plus honorary flatmates Farzaneh and Xiaowei have made Berkeley such a fun place to hang out or have deep philosophical discussions. Although many of us have scattered across the globe I look forward to meeting everyone soon.

I'd like to thank my good friend Daniel Keese for his friendship since 5th grade. I've been fortunate to have him an hour's drive away, and as he always reminds me, in the unlikely event that I'm in a position to hire an in–house lawyer, he'll be the first to know. I've also enjoyed hanging out with my college friend Faraz Ebrahimi and the fun times we've had around the bay.

I'd like to thank my friends Matthew Shuman and Dr. Jacob Postman for their friendship and all the good times we've had over the years. I've also had great times and helpful technical discussions with Mike Lorek, Dario Beraldo, Matthew Spencer, Milos Jorgovanovic, Matthew Weiner, Henry Barrow, Bobby Schneider, and too many others to mention. I'm truly thankful for all of the great people I've met here in Berkeley and in my travels around the world.

Even a blind chipmunk finds a nut in the woods every once in awhile, and I am so fortunate to have found Sasha. She is legitimately the awesomest (and most bad–ass) person I have met and I thank her for her love, friendship, and patience, and especially for her sense of adventure. I look forward to our next adventure.

Chapter 1

Introduction

Thanks to Moore's Law and the tireless efforts of three generations of engineers, the computer is now ubiquitous. Computers now assist or control almost every part of the daily lives of the citizens of the first world. The emergence of tablets and smartphones as primary interfaces to the internet has enabled enormous growth of the internet into the second– and third–world economies, reaching over a third of the world's population in 2013.

With this explosive growth has come a push to make user interfaces more natural. Already, user interfaces are undergoing a revolution. The touchscreen has largely replaced the traditional mouse and keyboard in mobile devices. Low cost inertial sensors enabled a new generation of hands-on gaming with great success. 3D imagers are now beginning to eliminate the controller from gaming consoles altogether. Gesture control is highly desirable for many additional applications, but unfortunately optical 3D imagers are too large, sunlight sensitive, and power hungry to be incorporated in an energy-constrained environment such as a mobile device.

The power dissipation of current optical depth sensors is on the order of one to ten watts. The power dissipation is dominated by the CMOS cameras and CMOS image processing chips which together form a depth image. Since incremental process improvements are unlikely to produce 2-3 orders of magnitude of power reduction before the next decade, a different approach is needed to build mobile-friendly gesture recognition systems.

Animals that live in environments that are not well-suited to optical vision rely on ultrasonic echolocation to navigate and hunt. We have built a system which emulates this evolutionary feature by building chip-scale ultrasonic depth sensors that enable gesture recognition with power dissipation of less than a milliwatt. We designed an energy efficient custom integrated circuit which interfaces with an array of aluminum nitride (AlN) piezoelectric micromachined ultrasound transducers (pMUTs) to transmit and receive ultrasonic pulses from the air.

Ultrasonic depth sensing enables gesture recognition in a small form factor with orders of magnitude lower power consumption than optical solutions. We are developing a design methodology for ultrasonic depth sensors which will allow energy efficient designs for a variety of applications. Simple, inexpensive MEMS fabrication combined with efficient CMOS interface circuits will ensure low cost, enabling wide adoption in mobile electronic devices.

This dissertation is divided into seven chapters. Following this introduction, the second chapter describes the operation of the micromachined ultrasonic transducer and how it can be modeled using simple circuit elements. The next chapter explains how time–of–flight rangefinding can be performed using a micromachined ultrasound transducer. The fourth chapter extends the approach to an array of transducers which permits the direction to the target to be estimated. The next chapter discusses the design of a energy–efficient readout circuit using a custom integrated circuit. The sixth chapter presents the results from this chip, and the final chapter offers conclusions and suggests avenues for future study.

Chapter 2

Piezoelectric Micromachined Ultrasound Transducers: Operation and Modeling

In this chapter, we discuss the operation of a piezoelectric micromachined ultrasound transducer (pMUT) which can be used to transmit and receive pressure waves from the air. We will show how the transducer can be modeled as a piston radiator and derive a circuit model for the transducer which accurately predicts the operation of the transducer. We will also discuss the use of a through–chip tube to improve acoustic coupling of the transducer to the air.

2.1 Basic Operation

The ultrasound transducer consists of a circular unimorph membrane consisting of a piezoelectric material sandwiched between two electrodes with an additional bending layer above or below the piezoelectric sandwich. We use AlN for the piezoelectric layer and AlN [1], SiO_2 [2], or Si for the bending layer. The device is fabricated on a Si wafer. As Figure 2.1 shows, a hole etched though the wafer exposes both sides of the membrane. The electrical field resulting from a voltage applied between the top and bottom electrodes results in a transverse stress in the AlN layer and consequent out-of-plane bending of the membrane, which produces a pressure wave. Similarly, an incident pressure wave results in membrane deformation and consequent charge on the electrodes, enabling the device to be used both as a transmitter and receiver. For optimum pMUT performance the top electrode should cover the area of the membrane where its displacement curve has positive curvature. For a membrane–style pMUT the optimum electrode/membrane diameter ratio size was found to be 70%. No dc bias is required for device operation.

Since the pMUT diameter/thickness ratio is very high, the residual stress plays an

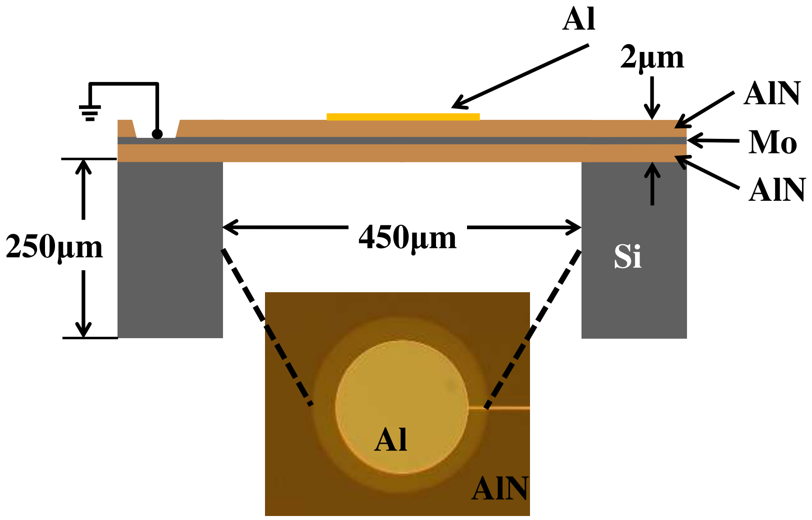


Figure 2.1: Schematic view of ultrasound transducer. *Inset:* Micrograph of single ultrasound transducer, showing partially transparent membrane.

important role in the device performance [3]. The pMUT displacement and resonance frequency are strongly dependent on the overall membrane residual stress. Another effect that can cause variations is surface roughness of the bending layer, which influences the quality of the piezoelectric AlN layer, and hence the device performance. The backside etch must stop on the membrane, so over-etch could also cause some variation in the displacement and resonant frequency of the device.

2.2 Basic Modeling of a Micromachined Transducer

The micromachined membrane can be modeled using standard resonator modeling techniques [4]. The membrane is a distributed mass–spring–damper, with the mass and stiffness primarily coming from the membrane material's mass and bending stiffness, while the damping comes primarily from the air. The transducer is used at resonance, at which point the energy stored in the membrane oscillates between stored (bending) potential energy and kinetic (inertial) energy. In order to calculate the modal mass, we need to find the average movement of the membrane during oscillation. This is called the modeshape of the transducer, and it can be approximated as:

$$d(r) \propto \left(1 - r^2\right)^2 . \tag{2.1}$$

This shows that the displacement of the transducer is maximum at the center and zero at the edges. For the purposes of calculating the mass of the transducer and the radiated sound, we want to calculate the effective surface area SA as if the transducer moved like a piston, with a uniform displacement.

We can integrate (2.1) to solve for the effective surface area SA of the membrane, and the result is that SA $\simeq 0.3\mathrm{SA}_{\mathrm{tot}}$. As a result, the effective radius of the device a_{ac} for acoustic calculations is

$$a_{ac} \simeq \frac{\sqrt{3}}{3} a. \tag{2.2}$$

Then, the modal mass is:

$$m_m \simeq 0.3\mathrm{SA}_{\mathrm{tot}} t_m \rho_m = \mathrm{SA} t_m \rho_m, \tag{2.3}$$

where t_m is the thickness of the membrane and ρ_m is the effective density of the membrane.

The zero stress resonant frequency of the membrane is given by [5]:

$$f_n \simeq \frac{10.18 t_m}{2\pi a^2} \sqrt{\frac{E}{12\left(1-\nu^2\right)\rho_m}}, \tag{2.4}$$

where a is the radius of the membrane, E is the effective Young's modulus of the membrane, and ν is the effective Poisson ratio of the material. From (2.4) and (2.3) and $\omega_n^2 = \frac{k_m}{m_m}$, we can say that the motional stiffness k_m' is

$$k_m' \simeq \frac{0.648 t_m^3}{\pi a^2} \frac{E}{\left(1-\nu^2\right)}. \tag{2.5}$$

In practice, the membrane has residual stress which can increase or decrease the effective stiffness of the device. Tensile stress serves to increase the resonant frequency, similarly to tightening a guitar string. Compressive stress reduces the resonant frequency. We can model the stress as an additional stiffness k_s, and so the total stiffness of the membrane is $k_m = k_m' + k_s$. Therefore the resonant frequency of the stressed transducer is

$$f_o \simeq \frac{1}{2\pi} \sqrt{\frac{k_m}{m_m}}. \tag{2.6}$$

A voltage applied to the piezoelectric membrane causes a force to develop across the membrane. The electrode configuration results in a parallel plate capacitance

$$C_o \simeq \frac{3\varepsilon_r SA}{2t_m}, \tag{2.7}$$

where the factor of $\frac{3}{2}$ comes from the fact that the electrode covers about $\frac{3}{2}$ of the effective surface area of the transducer.

In the mechanical domain, force F is modeled as a scalar quantity, and velocity v_m as a vector quantity. The electromechanical coupling factor η relates the voltage applied to the membrane to the force exerted by the membrane as $F = \eta V$, and the current contributing to the mechanical velocity v_m as $i_m = \eta v_m$ The electromechanical coupling factor is [2]:

$$\eta \simeq 2\pi(1.427)e_{31f}z_p, \tag{2.8}$$

where e_{31f} is the thin–film piezoelectric coefficient for the piezoelectric material, and z_p is the distance of the piezoelectric layer from the neutral axis of the membrane. In a properly designed transducer, z_p is nearly proportional to the thickness of the piezoelectric layer.

2.3 Acoustic Radiation from a Baffled Source

The air damping of the transducer due to acoustic radiation into the air is the dominant damping experienced by the transducer. The output force is converted to an output pressure by dividing by the surface area SA. In the acoustical domain, the pressure p and the volume velocity $V_v = v_m\mathrm{SA}$ describe the output acoustic power as $P_{ac} = pV_v$. The acoustic impedance D relates the pressure to the volume velocity as $p = DV_v$.

In this section, we first derive the bandwidth, transducer model, and output pressure for a membrane transducer with a baffle. In the following section we show how the addition of a through–wafer acoustic resonator tube can enhance the output pressure and bandwidth compared to a conventional baffled transducer.

Following [6], we can write the acoustic impedance D_{piston} experienced by a baffled piston as:

$$D_{piston}(a) = \frac{\rho c}{\pi a^2}\left(1 - \frac{2J_1(4\pi\frac{a}{\lambda})}{4\pi\frac{a}{\lambda}} + j\frac{2K_1(4\pi\frac{a}{\lambda})}{4\pi\frac{a}{\lambda}}\right), \tag{2.9}$$

where a is the effective radius of the piston. D_{piston} is called the piston function,[1] which describes the radiation efficiency of a circular piston. Figure 2.2 shows the normalized radiation impedance $D_{norm} = \frac{\mathrm{SA}}{\rho c}D_{piston}$ extracted from (2.9).

To aid in design, we can simplify D_{piston} by using a asymptotic circuit model which captures the behaviour of D_{piston} sans the oscillatory high frequency behavior. Figure 2.3 shows the resulting equivalent circuit. The resulting values for the circuit elements are [7]:

$$R_1(a) = 0.441\frac{\rho c}{\pi a^2}, \tag{2.10}$$

$$R_2(a) = \frac{\rho c}{\pi a^2}, \tag{2.11}$$

---

[1] J_1 is the Bessel function of order 1, and K_1 is the Struve function of order 1.

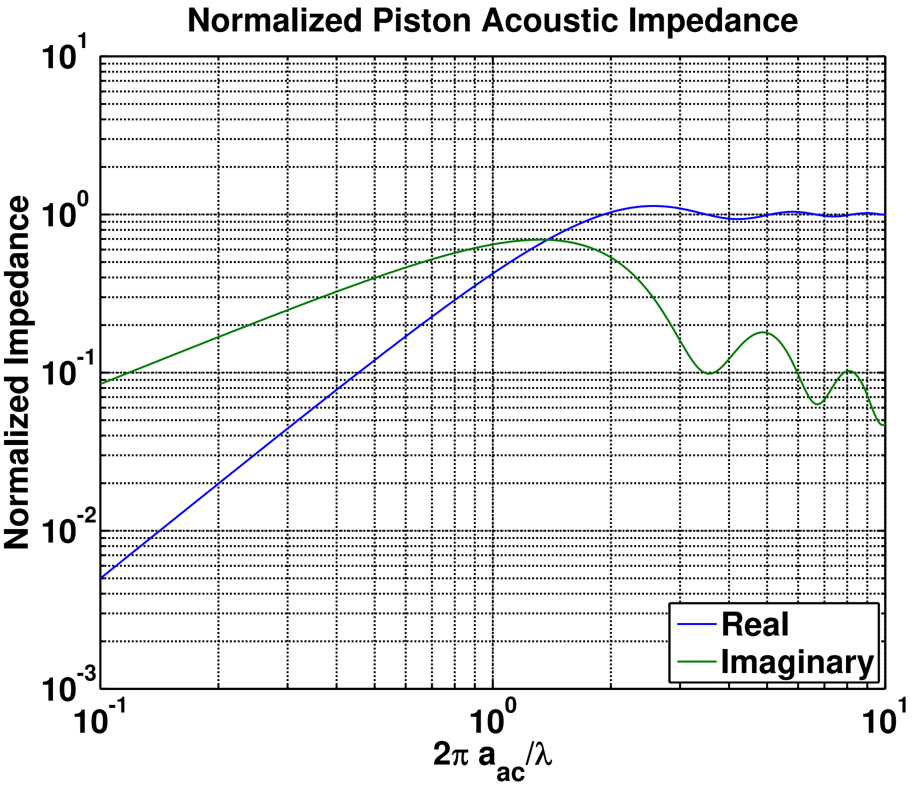


Figure 2.2: Normalized radiation impedance from a baffled piston source.

$$C_1(a) = 5.94\frac{a^3}{\rho c^2}, \tag{2.12}$$

$$L_1(a) = 0.27\frac{\rho}{a}. \tag{2.13}$$

Here a is the effective acoustic radius of the piston. Figure 2.4 shows the asymptotic model together with D_{piston}.

Since the edges of the ultrasound transducer do not move as much as the center does, the modeshape in (2.1) results in a reduced effective acoustic radius $a_{ac} \simeq \sqrt{0.3}a$. The use of this acoustic radius permits use of (2.9) to approximately describe the radiation behavior of the membrane without the use of complicated integrals of the 2D modeshape of the device to derive the device behavior.

Operation at frequencies below $\frac{a_{ac}}{\lambda} \simeq \frac{1}{\pi}$ results in reduced coupling to the air, reducing the bandwidth of the device; nevertheless, for a micromachined transducer, the device will usually be operated in this range. The damping of the device is $b_m = 2\mathrm{SA}^2\mathrm{Re}(D_{piston})$, where the factor of 2 comes because the device is assumed to be surrounded by air on both sides. The imaginary part of D_{piston} also contributes an

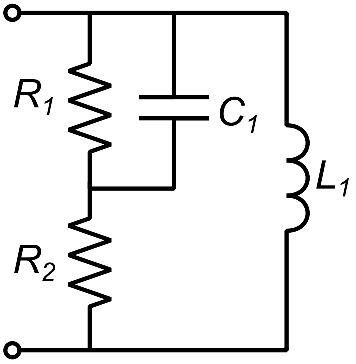


Figure 2.3: Asymptotic equivalent circuit model for D_{piston}.

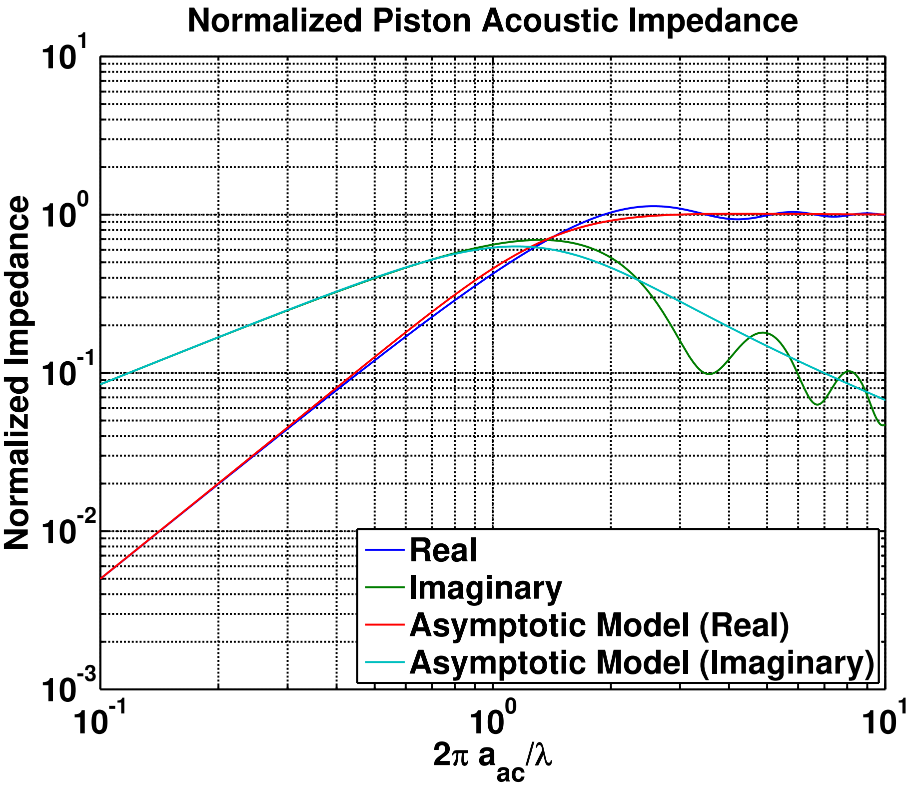


Figure 2.4: Normalized radiation impedance from a baffled piston source, with asymptotic equivalent circuit model.

additional mass $m_a = \frac{\mathrm{SA}^2\mathrm{Im}(D_{piston})}{2\pi f}$, which is usually small compared to the membrane mass m_m and can be neglected to the first order. To simplify the expression for (2.1), we can use the Maclaurin expansion of the real part of D_{piston}, and then $\mathrm{Re}(D_{piston}) \simeq$

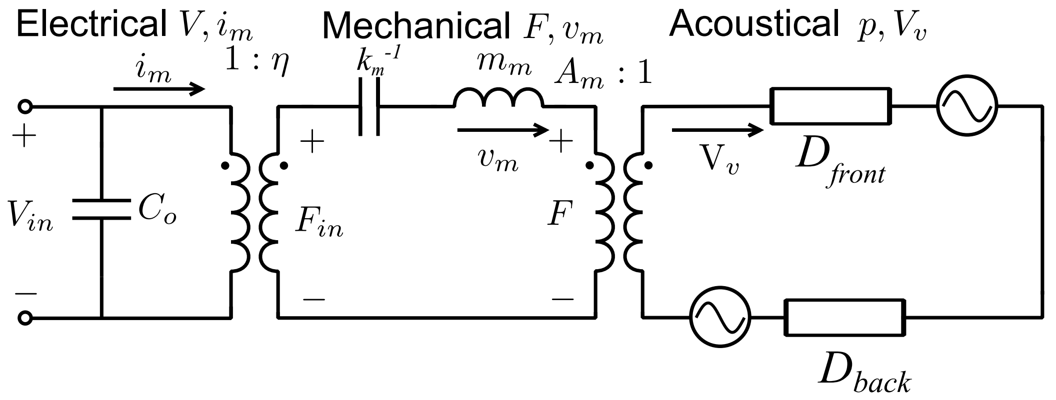


Figure 2.5: Equivalent circuit model for the baffled ultrasonic transducer which includes the electrical, mechanical, and acoustical domains.

$\rho c\frac{2\pi}{\lambda^2}$ and then the mechanical damping b_m is:

$$b_m \simeq 2\rho\frac{\mathrm{SA}^2\omega_o^2}{2\pi c} = \rho\frac{4\pi^3 a_{ac}^4 f_o^2}{c}. \tag{2.14}$$

The bandwidth of the transducer is

$$\mathrm{BW} = \frac{1}{2\pi}\frac{b_m}{m_m} \simeq 2c\frac{\rho}{\rho_m}\frac{\pi a_{ac}^2}{t_m\lambda^2} = 2\frac{\rho}{\rho_m}\frac{\pi a_{ac}^2 f_o^2}{ct_m}. \tag{2.15}$$

As explained in Chapter 2, large bandwidth is desirable for increased ranging accuracy and short minimum range. As (2.15) shows, in order to maximize bandwidth we would like to maximize the effective radius a_{ac} relative to the wavelength λ, while minimizing the thickness and density of the membrane. For reasons discussed in Chapter 4, the wavelength is usually chosen due to the desired application, where bandwidth is only one of many considerations, so in order to maximize the bandwidth we need to maximize the radius and minimize the thickness without reducing the operating frequency of the transducer. This means that the acoustic velocity of the membrane material is critical to large bandwidth. Aluminum nitride has an acoustic velocity of over 10km/s which compares favorably to many other materials, including PZT which has an acoustic velocity of only 4km/s. The use of a high velocity material such as CVD diamond could further increase the bandwidth of acoustic devices which are smaller than the wavelength of sound and which do not use coupling structures.

Figure 2.5 shows an equivalent circuit model for a simply baffled ultrasonic transducer which includes the electrical, mechanical, and acoustical domains.

The acoustic output pressure can be calculated from the equivalent circuit model. At resonance, the membrane dynamics cancel each other out and the force due to the piezoelectric moment is seen across the acoustic impedances. Following Blackstock [6] and Rayleigh, we calculate the output pressure that would exist at the face of the transducer if the piston diameter were infinite. This fictitious presssure p_o is used in the Rayleigh propagation equation presented in Chapter 3. Then the fictitious output pressure $p_o = \rho c v_m$ due to an input voltage V_{in} is given by

$$p_o = \rho c \frac{\eta V_{in}}{b_m} \simeq c^2 \frac{\eta V_{in}}{4\pi^3 a_{ac}^4 f_o^2}. \tag{2.16}$$

This equation together with (2.15) clearly shows the tradeoff between output pressure and bandwidth.

2.4 Use of an Acoustic Resonator to Improve Output Coupling

A resonant acoustic tube can be used to enhance the output pressure and the bandwidth of the transducer. The basic idea is to design a tube such that the wave reflected from the end of the tube interferes constructively with the output wave from the transducer. In this way a standing wave can be built up in the acoustic cavity.

The acoustic impedance inside a cylindrical tube with radius a is [6]:

$$D_{tube} = \frac{\rho c}{\pi a^2} \tag{2.17}$$

The wave travels down the tube at c, and so the time delay of a tube with length l is l/c. Neglecting loss from the walls and air absorption, the acoustic tube can be modeled as a transmission line with characteristic impedance D_{tube} and characteristic delay l/c. Figure 2.6 shows how we add the transmission line between the acoustic loads D_{front} and D_{back} which now model the frontside radiation (from the top surface of the membrane) and the bottom side radiation (from the throat of the tube on the bottom of the chip).

If the throat end of the tube (on the bottom of the chip) has the same radius a as the transducer, the air moving in the tube will encounter a reduced acoustic impedance $D_{back} = D_{piston}(a)$ at the interface between the tube and the bottom of the chip according to (2.9). The reflection ratio at the throat of the tube is then

$$R_{th} = \frac{D_{back} - D_{tube}}{D_{back} + D_{tube}}. \tag{2.18}$$

When a wave travels back down the tube towards the transducer, it will also reflect when it reaches the transducer. The reflection at the transducer is described by the acoustic impedance looking into the transducer membrane, which is the mechanical

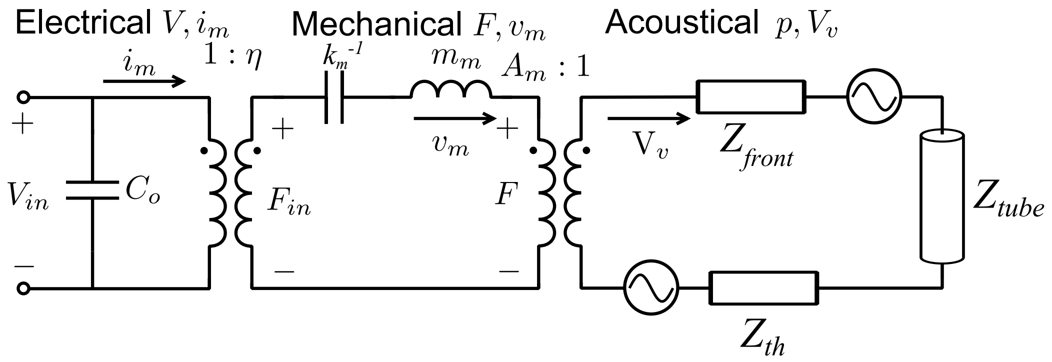


Figure 2.6: Equivalent circuit model for the ultrasonic transducer with a acoustic resonator coupling tube. The model includes the electrical, mechanical, and acoustical domains.

mass and stiffness, transformed into the acoustic domain, plus the acoustic damping from the frontside:

$$D_m \simeq \frac{1}{\mathrm{SA}^2}\left(j\omega m_m + \frac{k_m}{j\omega}\right) + D_{piston}(a_{ac}). \tag{2.19}$$

Here we've neglected the effect of the transducer electrical capacitance C_o since $C_o \gg C_m$, the motional capacitance.

The reflection ratio at the membrane end of the tube is then

$$R_m = \frac{D_m - D_{tube}}{D_m + D_{tube}}. \tag{2.20}$$

We are most interested in how the tube affects the impedance presented to the transducer, the output sensitivity, and the input sensitivity. The impedance presented to the bottom of the transducer is important because it determines the bandwidth of the transducer as we saw in (2.15). The impedance presented to the bottom of the transducer can be found by finding the ratio of the pressure at the bottom of the transducer to the volume velocity. Using the transmission line approach, we can write the pressure at a distance x from the throat of the tube as:

$$P = P^+ e^{jkx} + P^- e^{-jkx}, \tag{2.21}$$

where $k = \frac{\omega}{c}$ is the wavenumber, and the negative exponent in the second term signifies a backwards traveling wave. Note that $P^- = RP^+$. The volume velocity at the same point in the tube is

$$V_v = \frac{P^+}{D_{tube}} e^{jkx} - \frac{P^-}{D_{tube}} e^{-jkx} = \frac{P^+}{D_{tube}}\left(e^{jkx} - R_{th} e^{-jkx}\right) \tag{2.22}$$

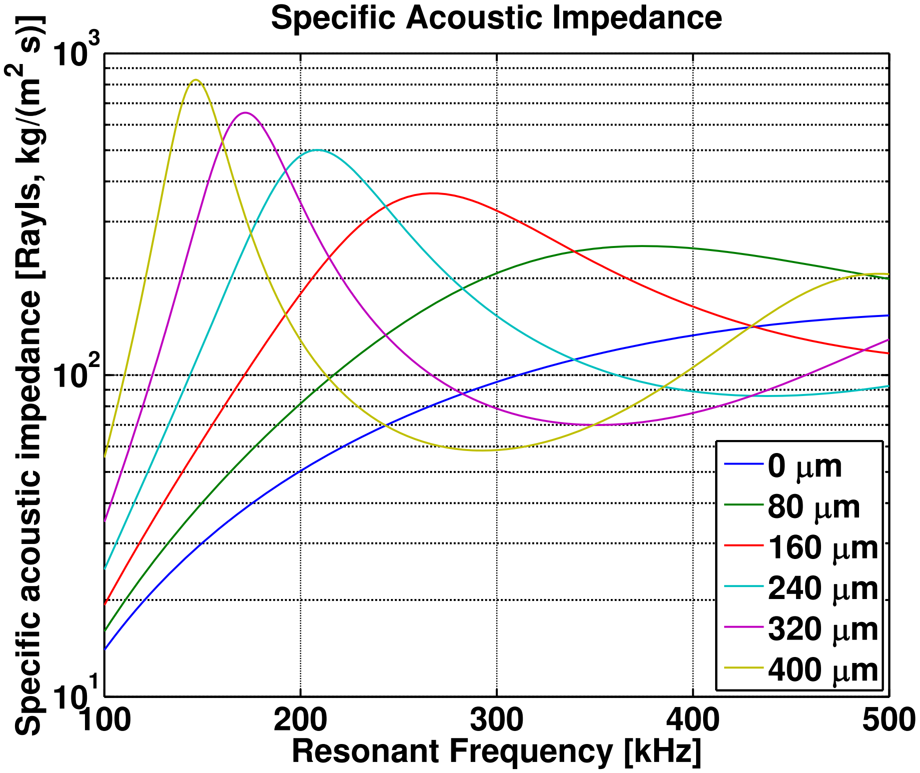


Figure 2.7: Specific acoustic impedance presented to transducer membrane at resonance for different lengths of tubes. Larger acoustic impedance results in larger bandwidth.

Therefore the familiar equation for impedance at any point in a transmission line also holds true for acoustic tubes:

$$D(x) = \frac{P^+ \left(e^{jkx} + R_{th} e^{-jkx}\right)}{\frac{P^+}{D_{tube}} \left(P^+ e^{jkx} - R_{th} e^{-jkx}\right)} = D_{tube} \frac{e^{jkx} + R_{th} e^{-jkx}}{e^{jkx} - R_{th} e^{-jkx}} \tag{2.23}$$

To find the acoustic impedance looking into the tube from the membrane, we simply plug the length of the tube l into this equation. We call the resulting input impedance of the tube D_b:

$$D_b = D_{tube} \frac{e^{jkl} + R_{th} e^{-jkl}}{e^{jkl} - R_{th} e^{-jkl}} \tag{2.24}$$

Figure 2.7 shows the resulting specific acoustic impedance presented to the membrane[2]. The effective impedance can be improved to larger than the ρc high frequency specific acoustic impedance of the air.

---

[2] specific acoustic impedance is in units of Rayls $\frac{\mathrm{kg}}{\mathrm{m^2 s}}$.

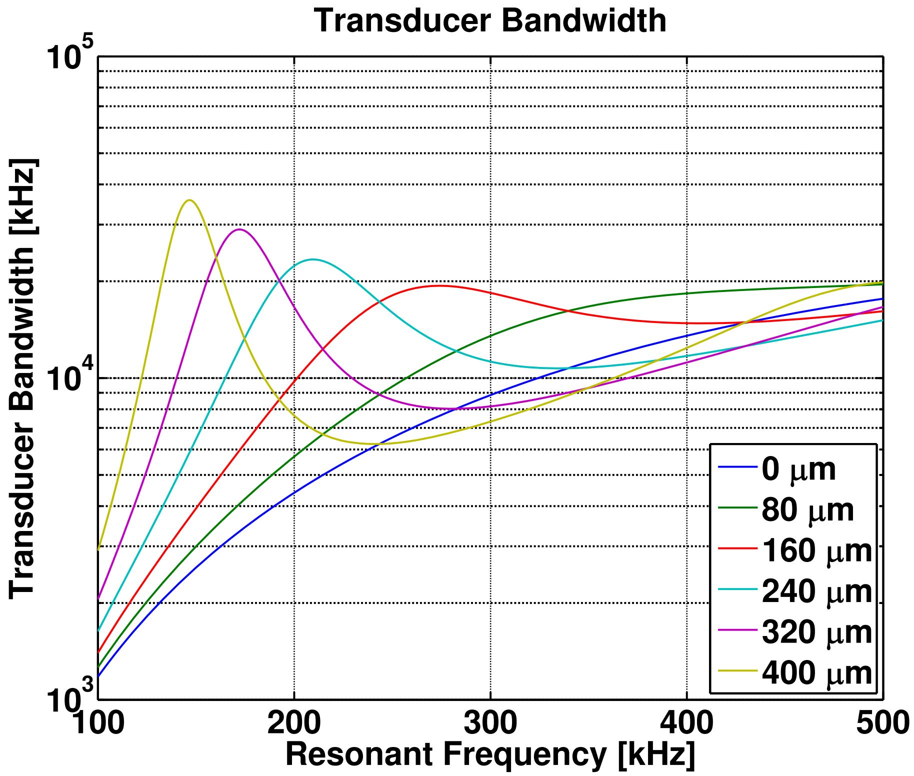


Figure 2.8: Bandwidth of transducer at resonance for different lengths of tube.

The increased acoustic impedance presented to the bottom of the transducer membrane results in an increased bandwidth; the bandwidth of the transducer is now:

$$\mathrm{BW} = \frac{1}{2\pi}\frac{b_m}{m_m} \simeq \frac{\mathrm{SA}^2\mathrm{Re}(D_{front} + D_b)}{2\pi m_m}. \tag{2.25}$$

To see how the transducer is affected by the addition of the tube, we can plot the bandwidth of a transducers of varying resonant frequency and tube lengths is plotted in Figure 2.8. In this plot, the radius of the transducer is held constant at 250μm, and the resonant frequency is adjusted by sweeping the stiffness of the transducer (similar to the effect residual stress would have on the transducer).

To find the transmit sensitivity, we use the Thevenin equivalent circuit shown in Figure 2.9. The equivalent acoustic output impedance at the throat D_{bin} is affected by the reflection at the membrane end of the tube, and so we can write it as:

$$D_{bin} = D_{tube}\frac{e^{jkl} + R_m e^{-jkl}}{e^{jkl} - R_m e^{-jkl}}. \tag{2.26}$$

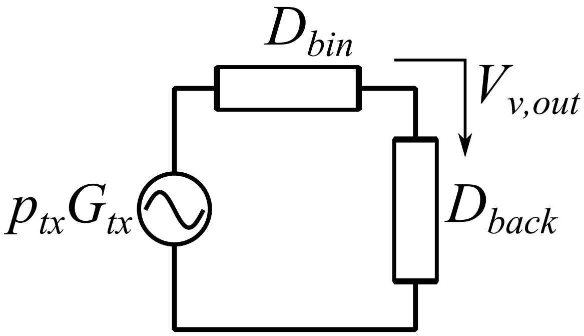


Figure 2.9: Thevenin equivalent circuit for determining output volume velocity from throat of tube.

Now at resonance, when the transducer is driven with a transmit voltage V_{TX}, the pressure across the acoustic terminal of the device is $p_{in} = \frac{\eta V_{TX}}{\mathrm{SA}}$. To calculate the equivalent pressure source shown in Figure 2.9, we need to include the gain from the acoustic terminal of the device to the open circuited output. This can be derived using a transmission line approach as was done above. The gain of the tube on the transmit side is

$$G_{tx} = \frac{-jD_{tube}\cos(kl)}{D_m\sin(kl) - jD_{tube}\cos^2(kl)} \tag{2.27}$$

Therefore the output volume velocity is

$$V_{v,out} = \frac{\eta V_{TX}}{\mathrm{SA}\,(D_{bin} + D_{back})}\frac{-jD_{tube}\cos(kl)}{D_m\sin(kl) - jD_{tube}\cos^2(kl)} \tag{2.28}$$

The tube benefits the output volume velocity as well as the bandwidth. Figure 2.10 shows the output volume velocity as a function of transducer frequency for differing lengths of tube. The plot is normalized to that of a transducer with the same radius but with no backside tube.

To calculate the input sensitivity of the device, we note that a received input pressure p_{rx} creates a force on the membrane which is converted to a voltage through the direct piezoelectric effect. We follow the same approach as above and use the Thevenin equivalent circuit shown in Figure 2.11. The output impedance of the circuit is $D_b + D_m$, which are shown in (2.24) and (2.19). The output equivalent pressure is calculated across D_{Co}, which is the impedance of the transducer's output capacitance C_o transformed into the acoustic domain:

$$D_{Co} = \frac{1}{j\omega C_o}\frac{\eta^2}{\mathrm{SA}^2}. \tag{2.29}$$

The acoustic gain due to the tube is:

$$G_{rx} = \frac{-jD_{tube}\cos(kl)}{D_{back}\sin(kl) - jD_{tube}\cos^2(kl)}. \tag{2.30}$$

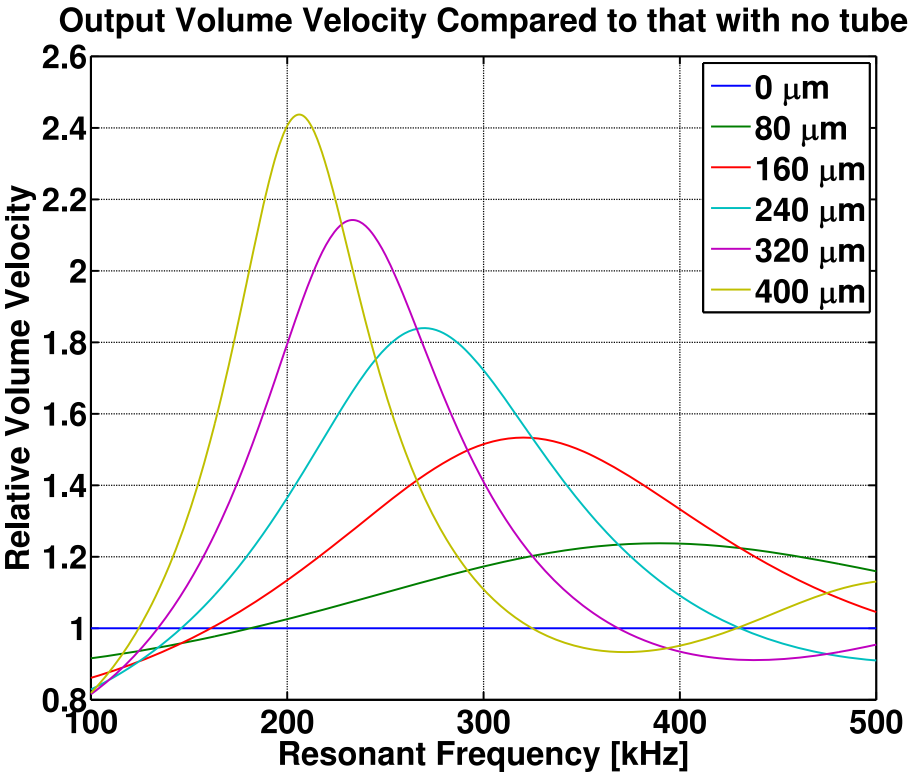


Figure 2.10: Volume velocity output of transducer for varying lengths of tube compared to the case when the tube is zero length and the transducer is simply baffled.

So the effective input voltage p_{in} is

$$p_{in} = p_{rx} G_{rx} \frac{D_{Co}}{D_{Co} + D_m + D_b}. \tag{2.31}$$

The input pressure is converted to a voltage $V_{in} = p_{in} \frac{\mathrm{SA}}{\eta}$.

The tube also increases the input pressure compared to what the input pressure would be with no tube. Figure 2.12 shows the relative increase in received signal compared to the case when no tube is used.

An acoustic resonator can significantly improve the bandwidth, output sensitivity, and input sensitivity of a transducer. As we will see in the next chapter, these three things are critical to building a rangefinder with good range and resolution.

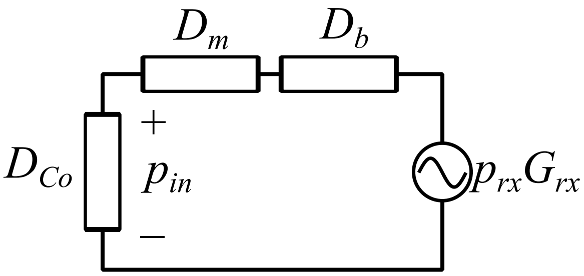


Figure 2.11: Thevenin equivalent circuit for determining input pressure at bottom side of membrane.

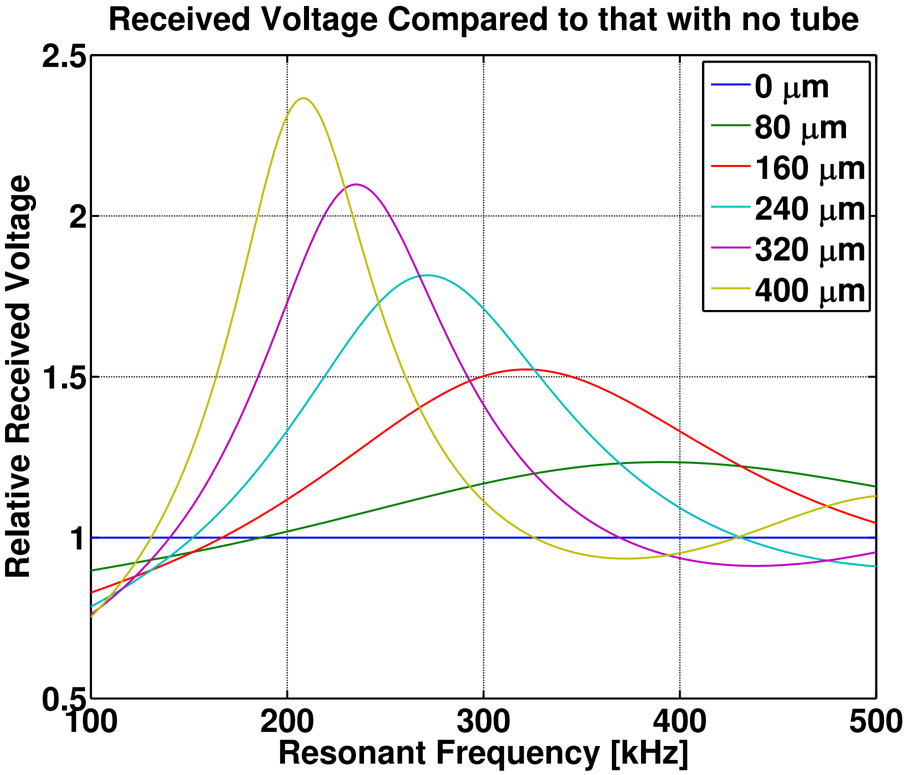


Figure 2.12: Voltage received by transducer for varying lengths of tube compared to the case when the tube is zero length and the transducer is simply baffled.

Chapter 3

Rangefinding with a Narrowband Ultrasound Transducer

In this chapter, we present the design equations for an ultrasonic rangefinder which measures the distance to a target in front of an aluminum nitride (AlN) piezoelectric micromachined ultrasound transducer (pMUT). We derive analytical expressions for the signal-to-noise ratio (SNR) and the random distance error, and use these determine how transducer parameters and material properties affect the performance of the system.

3.1 Rangefinding Basics

Rangefinding is a basic system-identification problem in which a sensor seeks to determine the distance to objects in the sensor's surroundings. Generally, the objects are within a direct line of sight to the sensor.

Time-of-Flight (ToF) measurements use the characteristic speed of a wave in a medium to determine the range to the object. In order to do so, the rangefinder must emit a signal into the medium and measure the response that returns. The system designer then has a choice between the use of electromagnetic radiation and mechanical oscillations of matter, or in other words, sound. In a ToF rangefinder, the rangefinder transmits a wave which travels at c, the speed of the wave in the medium, and returns after $T = \frac{2R}{c}$. In air, an electromagnetic wave travels at $c_l \simeq 3 \times 10^8 m/s$, the speed of light. On the other hand, sound travels at $c_s \simeq 343m/s$. So the reflection from an object at 1m returns after 6.6ns for electromagnetic excitation and after 5.8ms for sonic excitation. This motivates the use of sound to perform rangefinding, since relatively low–speed, low–power electronics can be used to send and receive signals with microsecond timing resolution, whereas achieving picosecond timing resolution requires high–speed electronics which require much higher power consumption.

Rangefinders operate either in continuous wave (CW) mode or pulse-echo (PE)

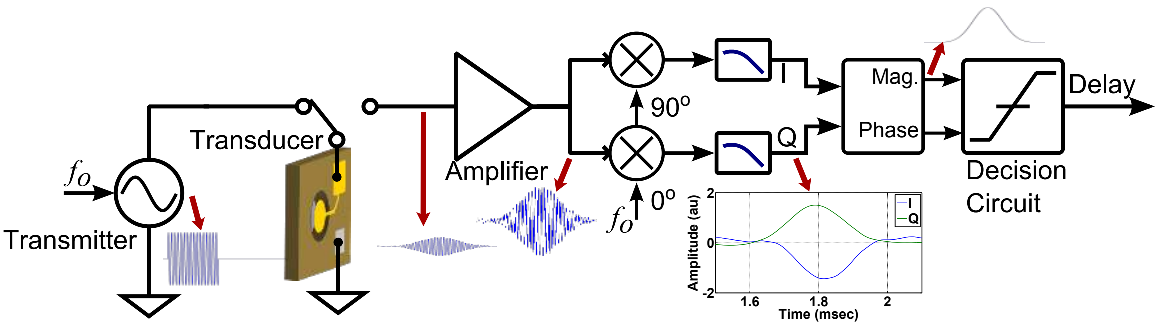


Figure 3.1: Block diagram of ultrasonic rangefinder.

mode. Narrowband CW systems suffer from multipath fading that can cause large range errors [8]. Frequency modulated continuous wave (FMCW) excitation can overcome multipath fading [9], but requires very high dynamic range since the transmitted signal dwarfs the return signal. PE excitation has lower average output power compared to CW, but the transmit pulse and return echoes are separated in time. This avoids the dynamic range and multipath problems that plague CW systems.

With these factors in mind, the remainder of this chapter details the system– and circuit–level considerations that are necessary when designing a pulse–echo ultrasonic rangefinder.

3.2 System Design

Figure 3.1 shows a block diagram of an ultrasonic rangefinder having a single transducer. For this system, the important design considerations are the transmit amplitude, transmit pulse shape, receiver noise, and the bandwidth of the receiver. These affect the maximum range and the distance error due to receiver noise.

Pulse transmission

A transmitter excites the transducer with an electrical pulse and causes the transducer to ring up and output a sonic pulse into the air. n ideal system would transmit a narrow, sharp pulse. Short duration permits resolving closely spaced targets, while short rise time is critical to minimize timing errors due to amplitude fluctuations. Consequently, in practical, band limited systems, it is desirable to use all of the available bandwidth. In practice, this means the bandwidth of the system should be designed to be limited by the bandwidth of the ultrasound transducer, and not the interface electronics.

The ideal transmit signal is approximated in practice with a burst of a sinusoidal signal at the resonant frequency of the transmitter element. The duration T_{TX} of the burst involves a tradeoff between minimum and maximum target range. Since the receiver is disabled during transmission, the minimum target distance is $R_{min} = \frac{c}{2}\mathrm{T}_{\mathrm{TX}}$. However, bursts much shorter than $T_{TX} = \frac{1}{BW}$ corresponding to $R_{min} \simeq \frac{c}{2BW}$ result in reduced transmit amplitude and hence smaller maximum range.

When operated in the linear regime at the resonant frequency, the transducer requires $\sim Q$ cycles to reach full output power. The output signal $p_{TX}(t)$ is

$$p_{tx}(t) = \frac{\eta V_{TX}}{\mathrm{SA}} G_{ac}\mathrm{u}(t)(1 - e^{-\omega_B t})cos(\omega_o t), \tag{3.1}$$

where η is the electromechanical coupling factor, V_{TX} is the transmit voltage, SA is the effective surface area, G_{ac} is the acoustic gain of transmitter, u(t) is the unit step function, ω_B is the radian bandwidth of the transducer, and ω_o is the radian resonant frequency.

When the transducer is operated in the nonlinear regime, the output power shows significant compression, but nonlinear operation still yields a substantial increase in the output power. Figure 3.2 shows a representative envelope of a transducer when it is excited in the nonlinear regime [10]. The onset of the nonlinear regime varies with different device design. Because of the compression, the transducer is less efficient at high output levels, and higher output levels lead to dimishing returns. As a result, each system has a lowest–energy optimum transmit level which depends on the maximum range and the efficiency of the receiver circuit. Depending on the design of the transducer, the maximum output pressure will either be limited by the maximum available transmit voltage or mechanical nonlinearity.

The design of the pulse shape is limited by the transducer characteristics. The narrowband nature of the transducer requires that ∼Q cycles are transmitted at the resonant frequency for the transducer to ring-up to full amplitude. If additional range is required, the duration of the transmit pulse can be increased, and the bandwidth of the bandpass filter following the front-end can be decreased, reducing the noise variance and thereby extending the range. However, as discussed next, this degrades the distance noise performance of the system. This tradeoff can be broken through the use of pulse compression [11], at the expense of added complexity to the system since in order to implement pulse compression the input signal must be correlated with the received signal.

The output pulse is shaped by the dynamics of the transducer. After the end of the transmit cycle, the stored mechanical energy in the transducer dissipates as the transducer rings down at the resonant frequency. Assuming the transducer is excited to full amplitude at f_o, the ringdown current $i_{ring}(t)$ is

$$i_{ring}(t) = \frac{V_{TX}}{R_M}\mathrm{u}(t - T_{TX})e^{-\omega_B(t-T_{TX})}cos(\omega_o t), \tag{3.2}$$

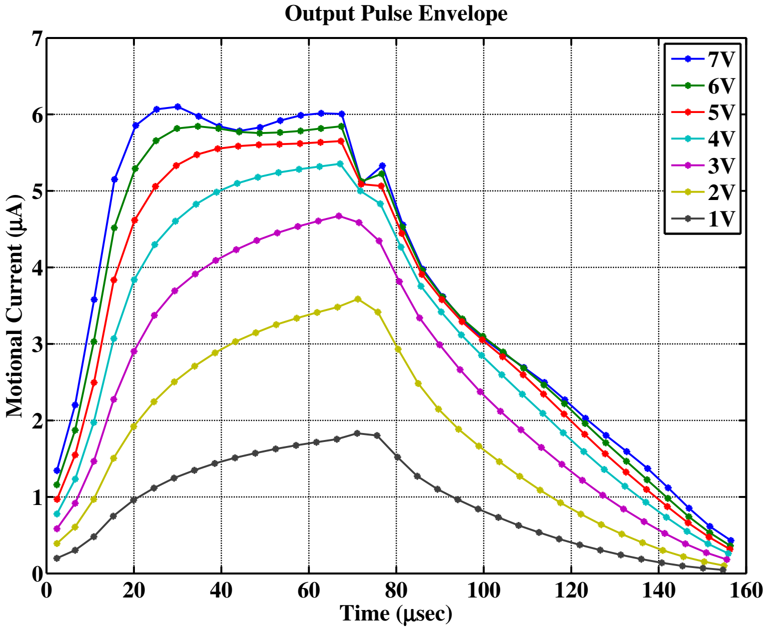


Figure 3.2: Measured output pulse shape for different drive voltages when the device is excited at f_o for $\sim Q$ cycles.

where T_{TX} is the transmit time and R_M is the motional resistance. In a system that uses a single element to send and receive, the ringdown signal can be recorded and subtracted from subsequent measurements to reduce the minimum range.

Pulse propagation

As the ultrasonic wave propagates away from the transducer, it spreads out and is also attenuated by thermal heating effects in the air. The path loss G_{path} is [6]:

$$G_{path} = G_{targ} \frac{r_r}{2R} 10^{-2\alpha R}, \tag{3.3}$$

where $r_r = \frac{\mathrm{SA_{eff}}}{\lambda}$ and α is the heating loss coefficient in Bels/m.

The heating loss is caused by three factors: (1) resonance of O_2 molecules in the air, (2) resonance of N_2 molecules in the air, and (3) thermoviscous damping. The resonance terms are affected by the presence of humidity in the air. At ultrasonic frequencies, the heating loss is dominated by thermoviscous damping. Figure 3.3 shows the heating loss absorption versus frequency. The heating loss coefficient α increases with f^2 in the high-frequency limit. As a result, increasing the operating frequency of a system reduces the maximum acheivable range.

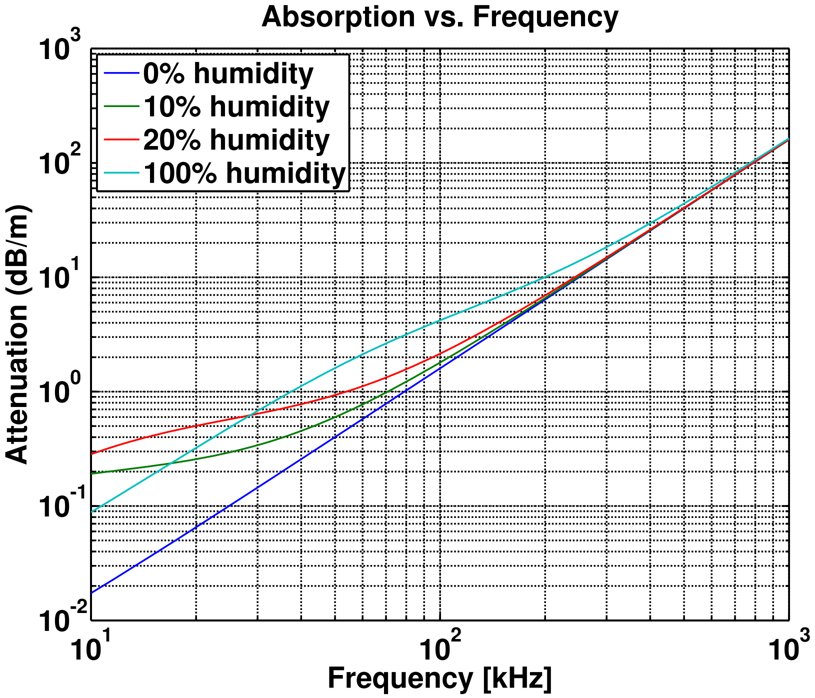


Figure 3.3: Heating loss absorption coefficient vs. frequency.

Figure 3.4 shows the resulting ultrasonic path loss versus range at 200kHz. The heating loss plays a relatively small role up to several hundred millimeters of range.

When the sound wave encounters an object, it is reflected in a direction normal to the surface of the target. For large targets, the object acts as a mirror and reflects the signal perfectly. For targets smaller than a wavelength, the object acts as a point source which re–radiates the reflected energy spherically, adding an additional $\frac{1}{R}$ term to the path loss equation. In general, many objects fall somewhere in–between these two regimes, and encounter an additional loss G_{targ} which is dependent on range and the area that is normal to the transducer.

Pulse reception

Returning echoes plus thermal motion of air molecules deflect the membrane and the echo is again shaped by the narrowband dynamics of the transducer. The transducer converts the mechanical velocity of the transducer into an output current. The front–end amplifier amplifies the signal and adds additional wideband thermal noise. The design of the front–end (FE) is therefore critical to the overall system design and will be presented in Section 3.3.

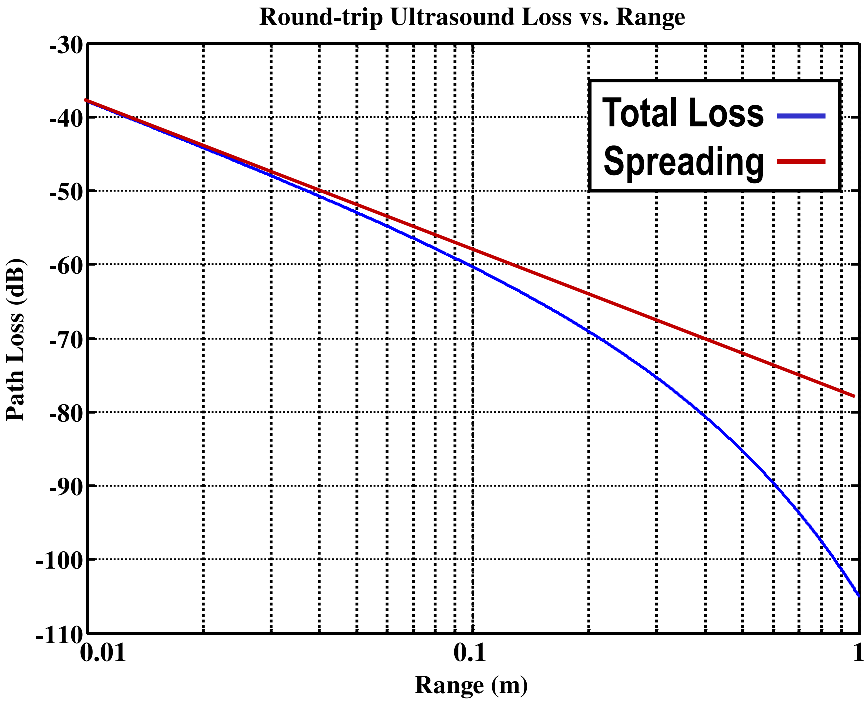


Figure 3.4: Ultrasonic path loss versus range at 200kHz.

The transmitted signal and the echo signal are illustrated in Figure 3.7. The desired signal is the time–of–flight. The receive signal is doubly shaped by the transducer dynamics; as a result the pulse shape is that of $p_{TX}(t)$ convolved with the impulse response of the transducer. The resulting pulse shape is best represented by a Gaussian window $S(t)$ whose width depends on T_{TX} and ω_B. The open circuit output voltage $V_{rx}(t)$ at resonance is

$$v_{rx}(t) \simeq p_{tx} G_{path} G_{ac} \frac{\mathrm{SA}}{\eta} \frac{1}{\omega_o C_o R_M} \mathrm{u}(t - \mathrm{ToF}) S(t - \mathrm{ToF}) cos(\omega_o t + \phi), \tag{3.4}$$

assuming $\omega_o C_o R_m >> 1$, where G_{path} is the acoustic path loss and ϕ is the phase shift of the signal with respect to the transmit signal due to the time–of–flight plus the phase shift due to the capacitance C_o, which is $\phi = \omega_o \mathrm{ToF} + arctan(\omega_o C_o R_m)$ on the interval $[0...2\pi]$.

Range Resolution

Targets spaced less than $\frac{c}{2}\mathrm{T_{TX}}$ in range result in overlapping echoes which are difficult to resolve as separate targets [11]. The pulse also broadens as it passes through the

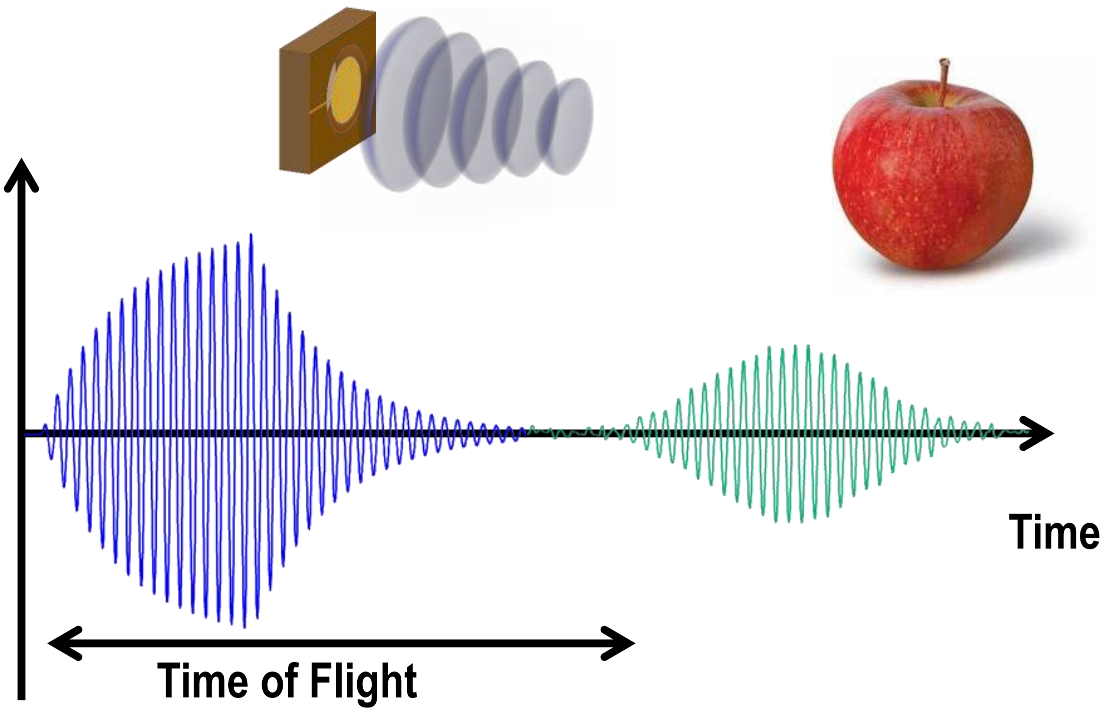


Figure 3.5: Illustration of output signal (blue) and received echo signal (green).

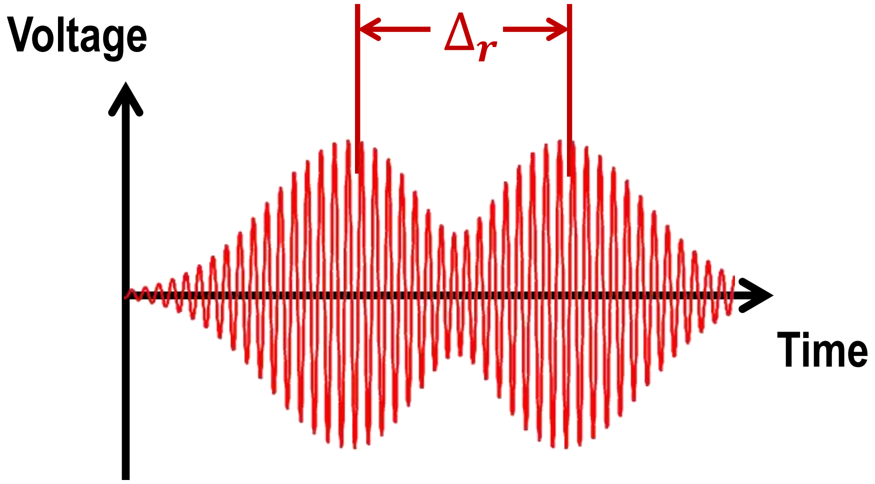


Figure 3.6: Echoes from closely spaced targets overlap, preventing the system from resolving them as separate targets.

narrowband transducer. For a pulse that is shaped by a second order transducer response on transmission and reception, the range resolution Δ_r is:

$$\Delta_r = \frac{c}{2}\left(T_{TX} + \frac{0.27}{\mathrm{BW}}\right). \tag{3.5}$$

For the minimum length pulse with full amplitude, and where $\mathrm{BW} = \frac{b_m}{2\pi m_m}$ and $T_{TX} \simeq \frac{1}{BW}$, (3.5) reduces to:

$$\Delta_r = \frac{c}{2}\left(\frac{1.27}{\mathrm{BW}}\right) = 1.27\pi c\frac{m_m}{b_m}. \tag{3.6}$$

This equation provides interesting insight about the influence of device scaling on bandwidth and hence resolution. When the transducer is large compared to the wavelength, or when a $\lambda/4$ matching tube is incorporated into the design as described in Chapter 1, then $b_m \simeq \rho c\mathrm{SA}$. Chapter 1 also shows that $m_m \simeq \rho_m t_m \mathrm{SA}$, where ρ_m and t_m are the density and thickness, respectively, of the transducer's membrane. Therefore,

$$\mathrm{BW} = \frac{\rho c}{2\pi\rho_m t_m} \tag{3.7}$$

$$\Delta_r = \frac{c}{2}\left(\frac{1.27}{\mathrm{BW}}\right) = 1.27\pi\frac{\rho_m}{\rho}t_m, \tag{3.8}$$

which shows that the transducer's bandwidth and resolution are independent of the speed of the medium, and simply dependent on the ratio of the density of the device to that of air and the thickness of the membrane.

Pulse Processing

Following the front–end amplification, the signal must be analyzed to extract the echoes of interest. The desired signal is the arrival time of the echo pulse which returns after a round–trip delay due to the speed of sound. A simplified signal processing scheme is shown in Figure 3.1. The signal must be filtered to remove wideband noise from the FE and demodulated to dc so that the decision circuit can extract echoes. Because the signal has an unknown phase, the demodulation must recover signals of any phase. A quadrature modulator is desireable since it preserves the signal phase information; simple systems could also use rectifying envelope detectors.

For pulse detection, a threshold is generally used to distinguish faint echoes from noise. The threshold must be set high enough to detect echoes at the maximum range but to reject noise pulses which would otherwise cause large errors in the range estimate. The average time between false alarms is given in [11]:

$$t_{fa} = \frac{1}{\mathrm{BW}}e^{\frac{V_{TH}^2}{\sigma_a^2}}, \tag{3.9}$$

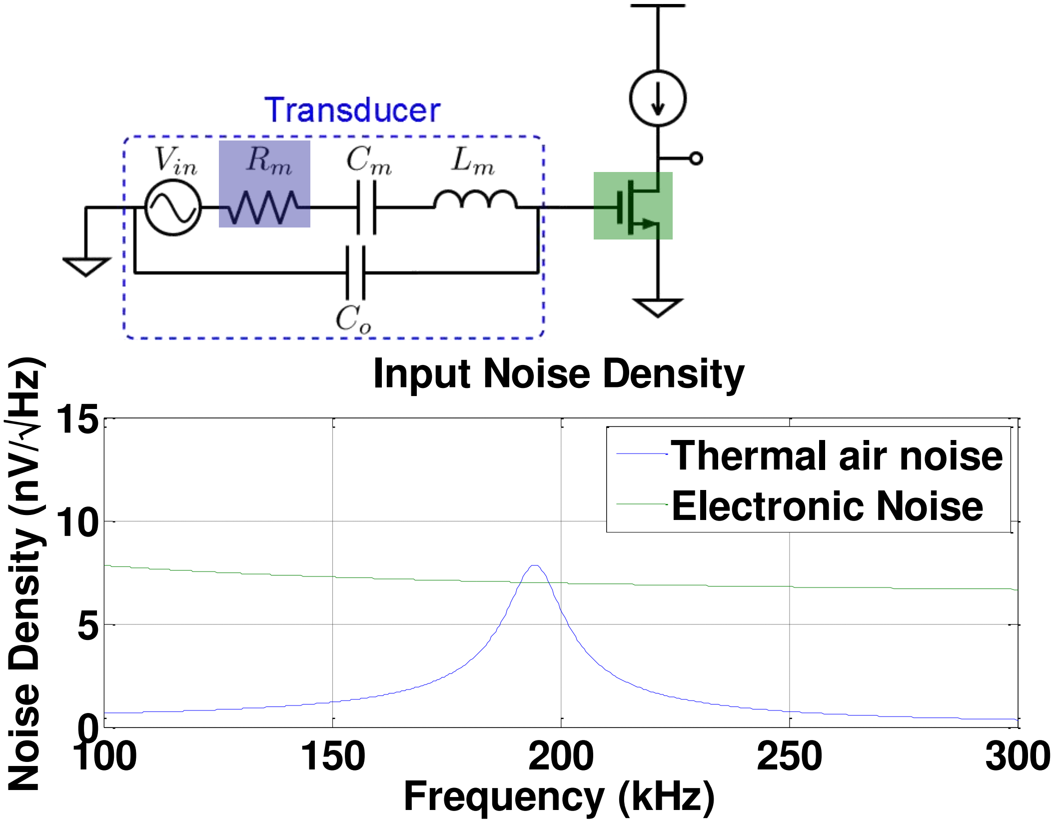


Figure 3.7: (top) Schematic showing important noise sources in the system. (bottom) Voltage noise density at the input of the amplifier.

where V_{TH} is the system's threshold and σ_a^2 is the amplitude noise variance of the system. The maximum range of the sensor is therefore determined by the minimum signal-to-noise ratio (SNR) required to reliably detect the received pulse and to reject false-alarms due to noise. Setting $t_{fa} = 30$ minutes results in $V_{TH}^2/\sigma_A^2 = 12$dB. Because of the exponent in (3.9), t_{fa} is very sensitive to V_{TH}^2/σ_A^2; reducing the margin to 11dB results in $t_{fa} = 54$ sec. Using $t_{fa} = 30$ minutes, the theoretical maximum range occurs when the SNR is equal to 12dB. The dominant noise sources in the system are the receive transducer and the front-end amplifier, discussed in more detail in the next section.

3.3 Noise analysis

There are two primary noise sources in a well–designed system: mechanical noise from the random thermal motion of the air that is received by the transducer, and electronic noise from the front–end amplifier.

Air molecules move in a gas at particle speeds described by the Maxwell-Boltzmann distribution, with an approximate average speed of the speed of sound:

$$c = \sqrt{\frac{\gamma k_b T}{m_{air}}}, \tag{3.10}$$

where $\gamma = 1.4$ is the adiabatic index of air[1], and $m_{air} = 4.93 \times 10^{-26}$kg is the average mass of an air molecule. The air molecules race around, undergoing brief elastic collisions with other nearby air molecules and solid surfaces. Collisions with the transducer membrane occur randomly and can be modeled as a white noise process. This white noise has a spectral density of $4k_bTR_m$. The mechanical noise is shaped by the transducer's mechanical dynamics and integrated onto the output capacitance of the pMUT. As a result, the integrated noise voltage $\overline{v_m}^2$ at the output of the transducer is

$$\overline{v_m}^2 \simeq \frac{k_b T}{L_M} \frac{1}{\omega_o^2 C_o^2}, \tag{3.11}$$

where k_b is Boltzmann's constant and L_M is the transducer's equivalent motional inductance.

The thermal noise added by the electronic amplifier can be modeled by analyzing a single MOSFET common source amplifier stage having transconductance g_m. The equivalent input noise voltage density is then $\frac{4k_bT}{g_m}$. The filter removes the out–of–band noise from the electronic amplifier. In the simplest case, we can assume that the noise bandwidth of the filter $\text{NBW} = \frac{2\pi \text{BW}}{4} = \frac{R_M}{4L_M} = \frac{\omega_B}{4}$. Adding the resulting integrated noise variance to (3.11), we have

$$\overline{v_n}^2 \simeq \frac{k_b T}{L_M} \frac{1}{\omega_o^2 C_o^2} + \frac{k_b T}{g_m} \frac{R_M}{L_M}, \tag{3.12}$$

which simplifies to

$$\overline{v_n}^2 \simeq \frac{k_b T}{L_M} \frac{1}{\omega_o^2 C_o^2} \left(1 + F\right), \tag{3.13}$$

where $F = \frac{R_M \omega_o^2 C_o^2}{g_m}$ is the noise penalty added by the amplifier.

---

[1] The adiabatic index accounts for the available rotational and vibrational modes of the molecule. Diatomic air molecules have energy in rotational modes but do not have significant energy in vibrational modes at room temperature.

3.4 Maximum Range

To calculate the maximum range, we need to determine the SNR of the input signal as a function of range to the target, and then use (3.9) to determine the minimum threshold needed to reject false alarms [12]. The rms value of the input signal given in (3.4) is

$$|v_{rx}| \simeq p_{tx} G_{path} G_{ac} \frac{\mathrm{SA}}{\eta} \frac{1}{\omega_o C_o R_M}, \tag{3.14}$$

where p_{tx} is the rms transmit pressure.

Then, combining (3.14) and (3.13), the signal–to–noise ratio (SNR) is

$$\mathrm{SNR} = \frac{|v_{rx}|^2}{\overline{v_n}^2} \simeq \frac{\left(p_{tx} G_{path} G_{ac} \frac{\mathrm{SA}}{\eta R_M}\right)^2}{\underbrace{\frac{k_b T}{L_M}}_{\text{Air Noise}} \underbrace{\left(1 + \frac{R_M \omega_o^2 C_o^2}{g_m}\right)}_{\text{Amplifier Noise Penalty}}}. \tag{3.15}$$

(3.15) can provide interesting insight about transducer design. First, let us re–write (3.15) in terms of transducer mechanical parameters by letting $\omega_o^2 = \frac{k_m}{m_m}$, $R_M = \frac{b_m}{\eta^2}$, $L_M = \frac{m_m}{\eta^2}$, and $G_{path}^2 = G_{targ}^2 \frac{\mathrm{SA}^2 \frac{k_m}{m_m}}{16 c^2 \pi^2 R^2} 10^{-4\alpha R}$. Then,

$$\mathrm{SNR} \simeq \underbrace{\frac{\frac{k_m}{m_m}\left(p_{tx} \frac{\mathrm{SA}^2}{b_m}\right)^2}{\frac{k_b T}{m_m}\left(1 + \frac{b_m k_m C_o^2}{\eta^2 m_m g_m}\right)} G_{ac}^2}_{\text{Design Parameters}} \underbrace{G_{targ}^2 \frac{1}{16 c^2 \pi^2 R^2} 10^{-4\alpha R}}_{\text{Physical Parameters}}. \tag{3.16}$$

In the case where p_{tx} is is limited by the available transmit voltage $\mathrm{V_{TX}}$, then $p_{tx} = u_o G_{ac} \rho c = \frac{\eta \mathrm{V_{TX}}}{b_m} G_{ac} \rho c$,

$$\mathrm{SNR} \simeq \frac{\frac{k_m}{m_m}\left(\eta \mathrm{V_{TX}} \frac{\mathrm{SA}^2}{b_m^2}\right)^2}{\frac{k_b T}{m_m}\left(1 + \frac{b_m k_m C_o^2}{\eta^2 m_m g_m}\right)} G_{ac}^4 G_{targ}^2 \frac{\rho^2}{16 \pi^2 R^2} 10^{-4\alpha R}. \tag{3.17}$$

Now, when the transducer is large compared to the wavelength, or when a $\lambda/4$ matching tube is incorporated into the design as described in Chapter 1, then $b_m \simeq \rho c \mathrm{SA}$. Plugging into (3.17), we have

$$\mathrm{SNR} \simeq \frac{\frac{k_m}{m_m}\left(\eta \mathrm{V_{TX}}\right)^2}{\frac{k_b T}{m_m}\left(1 + \frac{\rho c \mathrm{SA} k_m C_o^2}{\eta^2 m_m g_m}\right)} G_{ac}^4 G_{targ}^2 \frac{1}{16 \rho^2 c^4 \pi^2 R^2} 10^{-4\alpha R}. \tag{3.18}$$

In the case where the amplifier noise dominates, we can write (3.18) as

$$\mathrm{SNR} \simeq \left(\frac{g_m \eta^4 \mathrm{V_{TX}}^2}{k_b T \mathrm{BW} C_o^2}\right) G_{ac}^4 G_{targ}^2 \frac{1}{32 \rho^2 c^4 \pi^3 R^2} 10^{-4\alpha R}. \tag{3.19}$$

We can observe the following things about (3.19) for the case where the amplifier noise dominates:

- That SNR is dependent on the ratio of $\frac{\eta^4}{C_o^2}$. η is proportional to the piezoelectric material's e_{31} piezoelectric coefficient, and C_o is proportional to the piezoelectric material's ε_{33} permittivity. Therefore, the material property $\frac{e_{31}^2}{\varepsilon_{33}}$ is critically important to SNR.

- That SNR is directly proportional to g_m. g_m is proportional to the power consumption of the FE amplifier. Therefore, a 2× increase in power consumption leads to a 2× increase in SNR. In the case where spreading dominates G_{path}, this results in a $\sqrt{2}\times$ increase in the maximum range; in the case where absorption dominates, increased power consumption does not have a significant effect on the maximum range.

- That SNR appears to tradeoff with bandwidth ω_B. This is a false tradeoff, because we have assumed that the transmit pulse has the bandwidth of the transducer. If higher SNR is required, a shorter pulse can be used, which, if unmodulated, will increase the minimum achievable range resolution. This tradeoff can also be broken using pulse compression, whereby a signal with time–varying phase is transmitted. The receiver correlates the received signal with the transmit signal, and careful pulse design enables resolution equal to that which can be achieved with a short pulse [11]. As a result, it is not desirable to design a transducer with reduced bandwidth just to increase the SNR; pulse compression can do the same thing without decreasing the range resolution of the system. In the case where pulse compression is used, we can approximate (3.19) as:

$$\mathrm{SNR} \simeq \left(\frac{g_m \eta^4 T_{TX} {\mathrm{V_{TX}}}^2}{k_b T C_o^2}\right) G_{ac}^4 G_{targ}^2 \frac{1}{16\rho^2 c^4 \pi^2 R^2} 10^{-4\alpha R}. \tag{3.20}$$

3.5 Distance Error Sources

Distance measurement errors result from systematic errors and random errors. Systematic errors mainly come from environmental factors and fixed (i.e. offset and gain) errors from the receive circuit. The error due to the variation of the speed of sound with temperature is 1.75mm/m/^oC .

The finite rise time of the received pulse also leads to a systematic gain error, since the time between the beginning of the receive pulse and the crossing of the threshold depends on the amplitude of the pulse. This can be calibrated out, or eliminated by referencing the threshold to the pulse amplitude. Ideally, two thresholds would be used; an absolute threshold, which is derived from (3.9), and a relative threshold, which is equal to half the measured pulse amplitude.

For a time-of-flight based system, distance noise is created when amplitude noise is converted to timing noise by the finite rise time of the transmit pulse. Therefore, the ideal transmit waveform has zero rise time. Since any real system is band-limited, the bandwidth of the system limits the maximum achievable noise performance. The rms distance error σ_R for a linear detector is given by [11]:

$$\sigma_R = \frac{c}{2} \frac{1}{\kappa \mathrm{BW} \sqrt{\mathrm{SNR}}}, \tag{3.21}$$

where κ is a scaling factor that depends on pulse shape. Since the amplitude noise variance σ_A^2 is proportional to the bandwidth, the distance noise variance σ_d^2 decreases linearly as the bandwidth increases. For this reason, the optimal noise performance is achieved when the bandwidths of the transmit pulse, the transmitter, and the receiver are greater than the transducer bandwidth, so that the bandwidth of the signal is limited by the transducer only.

We can combine (3.17), (3.7), and (3.21) to find how the distance noise depends on device parameters:

$$\sigma_R^2 \simeq \frac{k_b T m_m \left(1 + \frac{b_m k_m C_o^2}{\eta^2 m_m g_m}\right)}{\underbrace{\kappa^2 \frac{k_m}{m_m} \left(\eta \mathrm{V}_{\mathrm{TX}} \frac{\mathrm{SA}^2}{b_m}\right)^2 G_{ac}^4}_{\text{Design Parameters}}} \frac{8c^2 \pi^3 R^2}{\underbrace{G_{targ}^2 \rho^2 10^{-4\alpha R}}_{\text{Physical Parameters}}}. \tag{3.22}$$

A pulse–echo rangefinding technique offers resistance to multi–path and dynamic range issues that plague continuous wave systems. This chapter showed that the key parameters that are important for rangefinding are BW and SNR, both of which should be maximized for good ranging performance. The next chapter extends the rangefinding technique to an array of rangefinders in order to measure the direction to the target in addition to the range.

Chapter 4

3D Rangefinding with an Array of Ultrasound Transducers

When more than one ultrasound transducer is used to send or receive sound from the air, the system can use the spatial arrangement of the transducers to infer the direction to the target. A 1D configuration of transducers permits a measurement of the angle to the target that is normal to the linear array; a 2D array permits the full 3D position of a target to be captured. This chapter presents design equations for 3D rangefinding with an array of ultrasound transducers.

4.1 Array Basics

The goal of an angle measurement system is to measure the direction to the target from the transceiver. Figure 4.1 depicts a linear array of receivers with neighboring elements spaced d distance apart. These can transmit a signal or receive a plane wave returning from a target. In the receive case, an plane wave returning at angle θ normal to the array creates a difference in the ToF of

$$\Delta T = \frac{d}{c} sin(\theta), \tag{4.1}$$

which corresponds to a phase shift

$$\Delta \phi = \frac{2\pi d}{\lambda} sin(\theta), \tag{4.2}$$

where λ is the wavelength of the sound wave. Similarly, to transmit a pulse with angle θ from normal, the phase of the transmit signal should be shifted by $\Delta\phi$ between each element.

For a linear array of N transducers spaced distance d apart, we define the aperture of the system as $A = \frac{Nd}{\lambda}$. Aperture in the angle axis is analogous to bandwidth in the

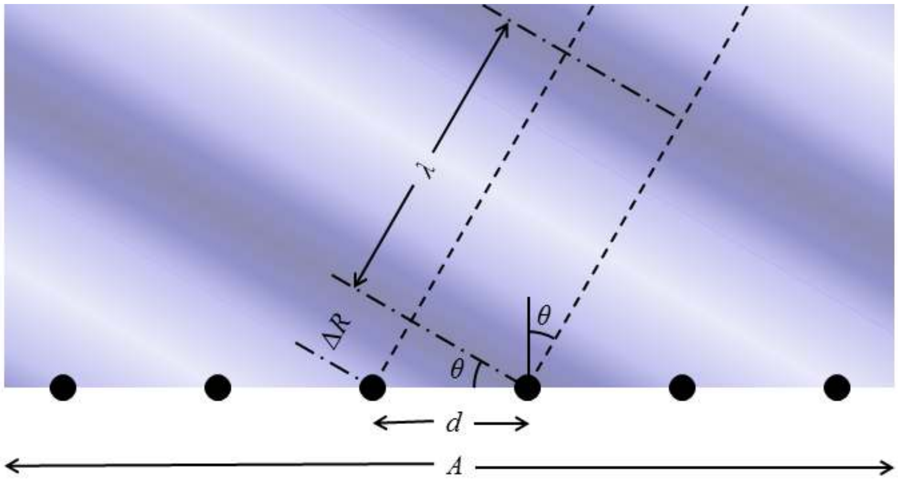


Figure 4.1: 1D ultrasonic receiver for measuring the angle θ of the target based on the difference in arrival time $\Delta\mathrm{T} = \Delta\mathrm{R}/\mathrm{c}$ between adjacent elements.

range axis, so increasing the effective aperture leads to increased angle measurement precision.

The aperture of the system determines the angular beam width of both the transmitted signal and the received signal. During transmission, the pressure waves are summed in the acoustic domain. During reception, the received signals can be summed in the electrical domain, either in an analog or digital fashion. An additional advantage of received summing is that the signals can be phase shifted by the circuit to align the received signals irrespective of the incoming angle. In either case, the directivity as a function of angle θ for a line array of isotropic sources can be shown [6] to be:

$$\mathrm{D}(\theta) = \frac{\sin\left(\pi \frac{Nd}{\lambda} \sin\theta\right)}{N \sin\left(\pi \frac{d}{\lambda} \sin\theta\right)}. \tag{4.3}$$

This expression is plotted for three cases of A in Figure 4.2. As $A = \frac{Nd}{\lambda}$ increases, the central lobe of the beam narrows. There are 3 ways to increase A: increase N, increase the spacing d, or decrease λ by increasing the operating frequency.

In the transmit case, the on–axis amplitude of the waves add, meaning the pressure on the axis goes with N. As a result, the power along the axis goes with N^2, which may be counterintuitive at first, since it seems to violate conservation of energy. However, the narrowing of the beam simply means that the power is concentrated along the axis, and the total radiated power still goes with N.

For the receive case, a similar result holds. Incoming signals that are in–phase can be summed and the resulting amplitude goes with N. Noise signals are normally

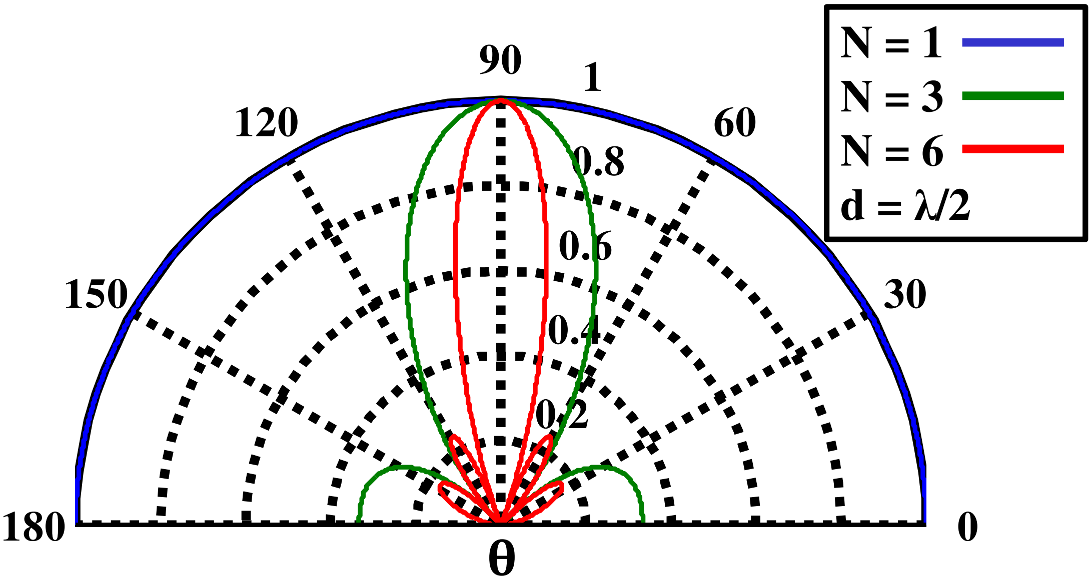


Figure 4.2: Directivity of a N element linear array with spacing $d = \lambda/2$.

uncorrelated between channels and therefore are added in a mean square fashion, and as a result the noise scales as $\sqrt{N}$.

For $N >> 1$, the small–angle approximation can be used in (4.3) to estimate the full–width half–power beamwidth θ_b as:

$$\theta_b \simeq \frac{50.76^o}{A}. \tag{4.4}$$

This equation gives the beamwidth for a single array. If transmit and receive arrays are used, the product of the transmit and receive directivity functions must be solved for θ_b. θ_b is the angular resolution of the system; for two distinct targets at the same range to be individually resolvable, they must be separated by at least θ_b angle.

The aperture also affects the angular accuracy of the system. The rms angle error σ_θ given in [11] is:

$$\sigma_\theta \simeq \frac{31.86^o}{A\sqrt{\mathrm{SNR}}}. \tag{4.5}$$

Since the SNR is maximum at short ranges, the angular accuracy can be several orders of magnitude lower than the angular resolution.

It may seem attractive to increase d to increase the aperture. As d increases above $\frac{\lambda}{2}$, A increases, and the central lobe narrows, but the sidelobes become more dramatic. The reason for this is that d is basically the spatial sampling period of the array. Imagine an incoming wave which is at 90^o from normal, or in other words parallel to the array. The wave is sampled at each of the array elements. In order to satisfy the Nyquist rate in the spatial domain, the wave must be sampled at least twice per period, or in other words $d \leq \frac{\lambda}{2}$. If this criterion is not met, aliasing occurs, and signals from large angles will be indistinguishable from those contained in the central lobe. We call the resulting lobes in the directivity pattern sidelobes and these appear increasingly large as d increases.

Signal to Noise Ratio enhancement

In addition to enabling 3D localization, an array of transducers can serve to increase the SNR of the rangefinder. Since the dominant noise sources in the system are in the receiver, increasing the number of receive elements increases the signal power by N^2 but also increases the noise power by N. On the other hand, adding transmit elements does not appreciably increase the total noise, but instead decreases the width of the transmit beam. Consequently, the on–axis SNR for a system with N_{tx} transmitters and N_{rx} receivers is

$$\mathrm{SNR} \simeq N_{rx} N_{tx}^2 \frac{\frac{k_m}{m_m}\left(\eta \mathrm{V}_{\mathrm{TX}} \frac{\mathrm{SA}^2}{b_m^2}\right)^2}{\frac{k_b T}{m_m}\left(1 + \frac{b_m k_m C_o^2}{\eta^2 m_m g_m}\right)} G_{ac}^4 G_{targ}^2 \frac{\rho^2}{16\pi^2 R^2} 10^{-4\alpha R}. \tag{4.6}$$

As shown in equation (4.6), an N-element receive array will only increase the SNR by a factor of N compared to a single transceiver. However, since the receive signals can be summed in the electrical domain, the signals can be shifted so that the phases align irrespective of the input phase difference. Because of this flexibility, a receive array can maintain a constant SNR across a wide span of angles.

A transmit beamformer, on the other hand, will reduce the width of the transmit beam, so in order to illuminate a wide area the beam must be swept across the environment. Conventionally, this is done by steering the transmit pulse along a certain axis, and listening for the echo that comes back. Then the procedure is repeated along several different directions until sufficient coverage of the environment is achieved. As a consequence, it takes approximately N_{tx} times as long to create a measurement of the entire scene when using transmit beamforming. Therefore, there is no advantage to using transmit beamforming as opposed to receive beamforming.

4.2 Beamforming in a 1D array

Beamforming is a technique that is used in radar, sonar, and medical ultrasound to combine transmit or receive signals such that they combine constructively along the desired direction. This is a form of spatial filtering which allows the system to maximize the received signal from a desired direction.

The basic idea is to delay signals such that when they arrive at the target, or are summed at the receiver, they are in phase with each other. For wideband systems, this requires a variable delay along each transmit/receive channel in order to match the delays of the signals. In narrowband systems, the delay between the signals is much shorter than the total transmit time, and therefore phase shifting is sufficient to align the signals during the majority of the pulse length.

Time delays are required when the signals at the extreme ends of the array do not overlap. (4.1) shows the time delay between each neighboring element; the delay between N elements must be less than the transmit time in order to use phase shifting. Therefore we can say the condition for narrowband beamforming to be feasible is:

$$T_{TX} > \Delta T = \frac{Nd}{c} sin(\theta). \tag{4.7}$$

In the worst case, $T_{TX} = \frac{1}{\mathrm{BW}}$ (from Chapter 3) and $\theta = 90^o$, then the condition becomes

$$A \times \mathrm{BW} < f_o, \tag{4.8}$$

or more elegantly

$$A < Q. \tag{4.9}$$

In many cases (4.9) will be true for an ultrasonic rangefinder in air, and usually also holds for radar, allowing use of narrowband beamforming. Underwater sonar and medical ultrasound generally require wideband time–delay beamformers.

(4.2) is the basic equation used in narrowband beamforming. For a system with N transducers spaced d apart, the phase shift of the wave arriving at element k relative to the phase at element 1 is

$$\phi_{1:k} = \frac{2\pi d(k-1)}{\lambda} sin(\theta), \tag{4.10}$$

Therefore to beamform a set of complex signals $s_i(t)$ from N transducers along angle θ_m we perform the following operation:

$$b_m(t) = \sum_{i=1}^{N} s_i(t) e^{j\pi\left(\frac{2\pi d(k-1)}{\lambda} sin(\theta)\right)}, \tag{4.11}$$

where j is the imaginary unit.

In a transmit beamforming system, the summation occurs in the acoustic domain naturally since pressure is a scalar field, and the pressures from each transducer add coherently. In the receive case, the summation is performed in the receiver electronics, usually in the digital domain.

Receive Beamforming

Receive beamforming can be performed across a wide range of angles, limited only by the directivity of the individual ultrasound transducers and the directivity of the transmitters. The idea is to perform the beamforming operation in (4.11) for a range of angles θ over which the target may be located. The interval between angles $\Delta\theta = \theta_m - \theta_{m-1}$ should be chosen to be less than half of the angular resolution θ_b in order to satisfy Nyquist. For a system which illuminates a full 180^o, the number of discrete angles M to perform beamforming at should be:

$$M = \frac{180^o}{\Delta\theta} \geq 2\frac{180^o}{\theta_b} \simeq 7.09A \tag{4.12}$$

The receiver can create a matrix from the resulting beamformed vectors $b_m(t)$ along a set of angles $\theta_m = [\theta_1 ... \theta_M]$:

$$\Theta(t,\theta) = \begin{bmatrix} b_1(t) & \cdots & b_M(t) \end{bmatrix}. \tag{4.13}$$

The resulting matrix, shown in Figure 4.3, depicts the intensity of the reflected signal vs. range and direction. This is similar to a b–mode (brightness mode) image from medical ultrasound.

An array of ultrasound transducers allows localization of one or more objects in front of the array. A 1D array permits 2D localization, and a 2D array permits 3D localization. The array can be used to create a transmit beam or a virtual beam in the receive electronics. The size of the array relative to the wavelength is called the aperture and this determines the angular resolution of the system. The aperture and the signal to noise ratio determine how accurately a single target can be localized.

In the next chapter, we will discuss how to design an ultrasonic rangefinder in a energy efficient manner using the concepts developed in this chapter and in the previous chapter.

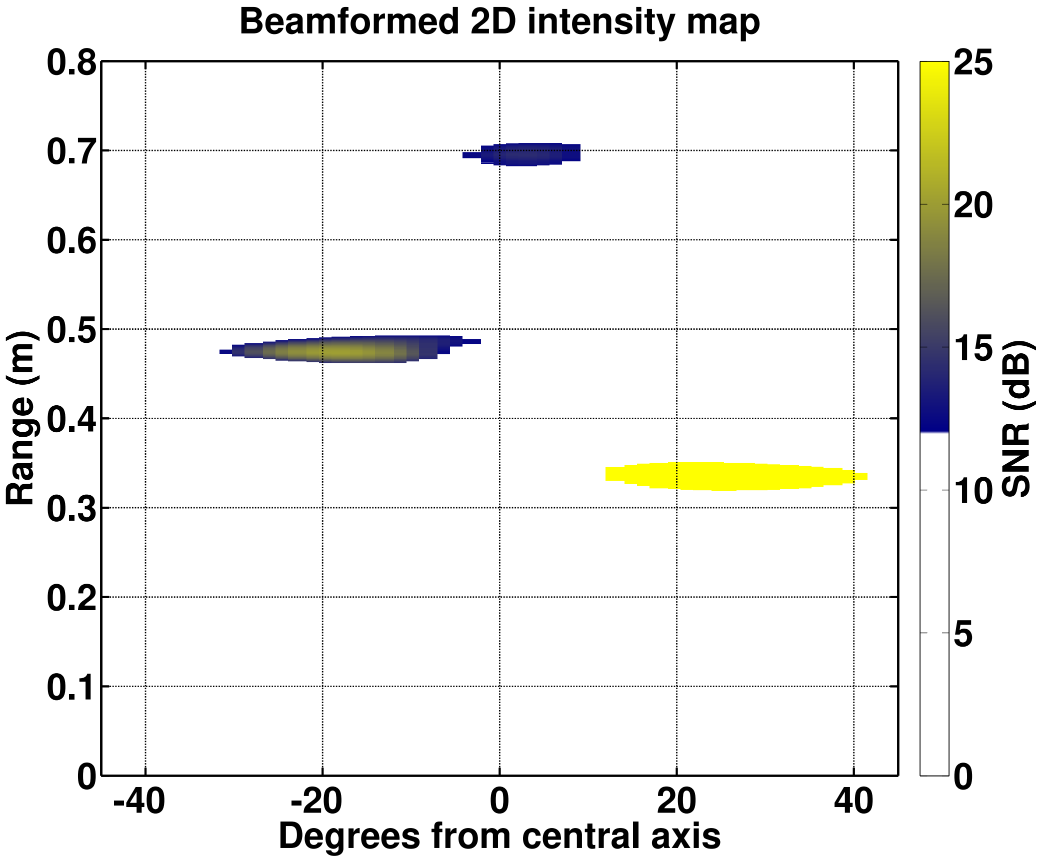


Figure 4.3: Beamformed 2D intensity map of environment.

Chapter 5

Energy Efficient Ultrasonic Rangefinder Readout Circuit Design

In this chapter, we present a methodology to design a energy efficient ultrasonic rangefinder interface circuit. We discuss the tradeoffs and main contributors to power consumption.

5.1 Rangefinder Block Diagram

In this section we develop a simple block diagram for the rangefinder system to motivate further discussion. We assume that we seek to design a system with multiple transducers which can be used to transmit or receive, and whose outputs should be combined in a digital beamformer as discussed in Chapter 4.

Since we would like each transducer to be capable of transmitting or receiving, each readout channel should have a transmitter and a receiver, plus a switch to isolate the receiver from the transmit pulse. The receiver should have a low–noise input and a digital output. The beamforming operation discussed in Chapter 4 requires a series of phase shifts to be implemented between the channels. This is best implemented in the digital domain since it is generally difficult or power hungry to generate accurate analog delays. Therefore, it is necessary to digitize each channel separately with an analog to digital converter (ADC) so that the signals can be combined in a digital beamformer.

The resulting block diagram is shown in Figure 5.1. In the following sections, we discuss each block design in detail.

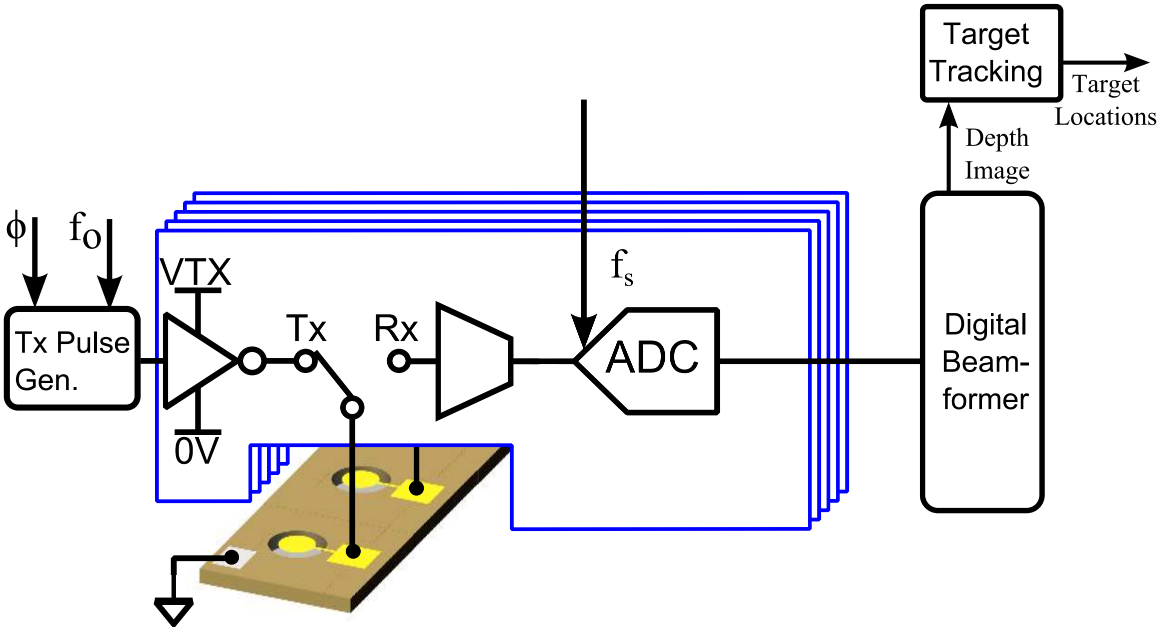


Figure 5.1: Simple block diagram illustrating the basic blocks required in a pulse–echo phased–array rangefinder.

Transmitter

The transmitter is used to drive the transducer at resonance to output a excitation pulse into the environment. The duty cycle of the transmitter is usually on the order of 1–2%, so a transmitter with no static power consumption or a power–down state is required for best efficiency. The narrowband transducer filters the output signal, and linearity usually is not an important consideration, so a nonlinear transmitter can be used for best power efficiency. In practice an inverter usually meets the requirements.

The energy required to drive each transducer is dominated by the energy required to charge and discharge capacitance of the transducer C_o. For a transmit voltage V_{TX}, each cycle $E_{cycle} = C_o V_{TX}^2$ energy is lost. As mentioned in Chapter 3, the minimum pulse that can be transmitted is approximately Q cycles in duration, so the transmit energy to drive N_{TX} transmitters is roughly $E_{TX} = N_{TX} \gamma Q C_o V_{TX}^2$, where γ is the transmitter efficiency, which can be degraded due to the transmitter's intrinsic capacitance and shoot–through current during switching.

The transducer has a range in which it responds linearly to excitation; outside this range the gain starts to compress. The compression varies due to device design, but a simple empirical model for the effective transmit voltage $V_{TX,eff}$ which includes the gain compression is:

$$V_{TX,eff} = \frac{V_{TX}}{1 + \frac{V_{TX}}{V_{crit}}}, \tag{5.1}$$

where V_{crit} is the critical input voltage where the transmit gain is reduced by a factor of 2. For a typical pMUT, $V_{crit} \simeq 16\mathrm{V_{rms}}$. For maximum gain the transducer should be actuated in a bipolar fashion. This can be achieved by switching the bottom electrode of the transducer to $V_{TX}/2$ during transmit.

Received Signal

The received signal consists of echo signals from targets in the environment, plus a ringdown signal that occurs as the transducer stops transmitting and settles into a static receive mode. Figure 5.2 shows that the received signal varies widely due to the range to the target, the size of the target, and the direction to the target. The echo signals contain information about the location of the targets in the environment encoded in the time–of–flight of the sound. The ringdown signal always occurs at the resonant frequency of the transducer f_o, irregardless of the transmit frequency, so measurement of the ringdown frequency can enable the electronic system to track the resonant frequency of the transducer.

Front–end Amplifier

The front–end amplifier serves to amplify or buffer the signal before the ADC converts it to digital form. The important characteristics of the front–end are the noise figure and the input and output impedance.

As explained in Chapter 3, the input voltage noise variance, including the contribution from the transducer and a common source amplifier stage having transductance g_m, is

$$\overline{v_n}^2 \simeq \frac{k_b T}{L_M}\frac{1}{\omega_o^2 C_o^2}\left(1 + \frac{R_M \omega_o^2 C_o^2}{g_m}\right) = \frac{k_b T}{L_M}\frac{1}{\omega_o^2 C_o^2}\left(1 + F\right). \tag{5.2}$$

The first term is the noise contributed by the transducer, and $F = \frac{R_M \omega_o^2 C_o^2}{g_m}$ is the noise penalty added by the amplifier. For a typical transducer, $R_M = 2.5\mathrm{M\Omega}$, $C_o = 10\mathrm{pF}$, and $\omega_o = 2\pi 200\mathrm{kHz}$, so $R_M \omega_o^2 C_o^2 \simeq 400\mu\mathrm{S}$. So if the amplifier has an effective g_m of $400\mu\mathrm{S}$, the noise figure $(1 + F)$ will be 3dB. This expression also holds for a operational transconducting amplifier (OTA) used with capacitive feedback, as long as the feedback capacitor $C_f \ll C_o$.

The power consumption of the front–end is determined by the desired noise factor since $g_m = \frac{2I_D}{V^*}$, where V^* is the short–channel effective overdrive voltage, and $P_{fe} = I_D V_{DD}$. We can then re–write the noise penalty as $F = \frac{R_M \omega_o^2 C_o^2}{2I_D} V^*$. The term $R_M \omega_o^2 C_o^2 V^*/2 \simeq 19.5\mu\mathrm{A}$ for $V^* = 0.1\mathrm{V}$, so a common–source amplifier with $19.5\mu\mathrm{A}$ current consumption will have a noise figure of 3dB. Figure 5.3 shows the theoretical in–band SNR following the front–end amplifier.

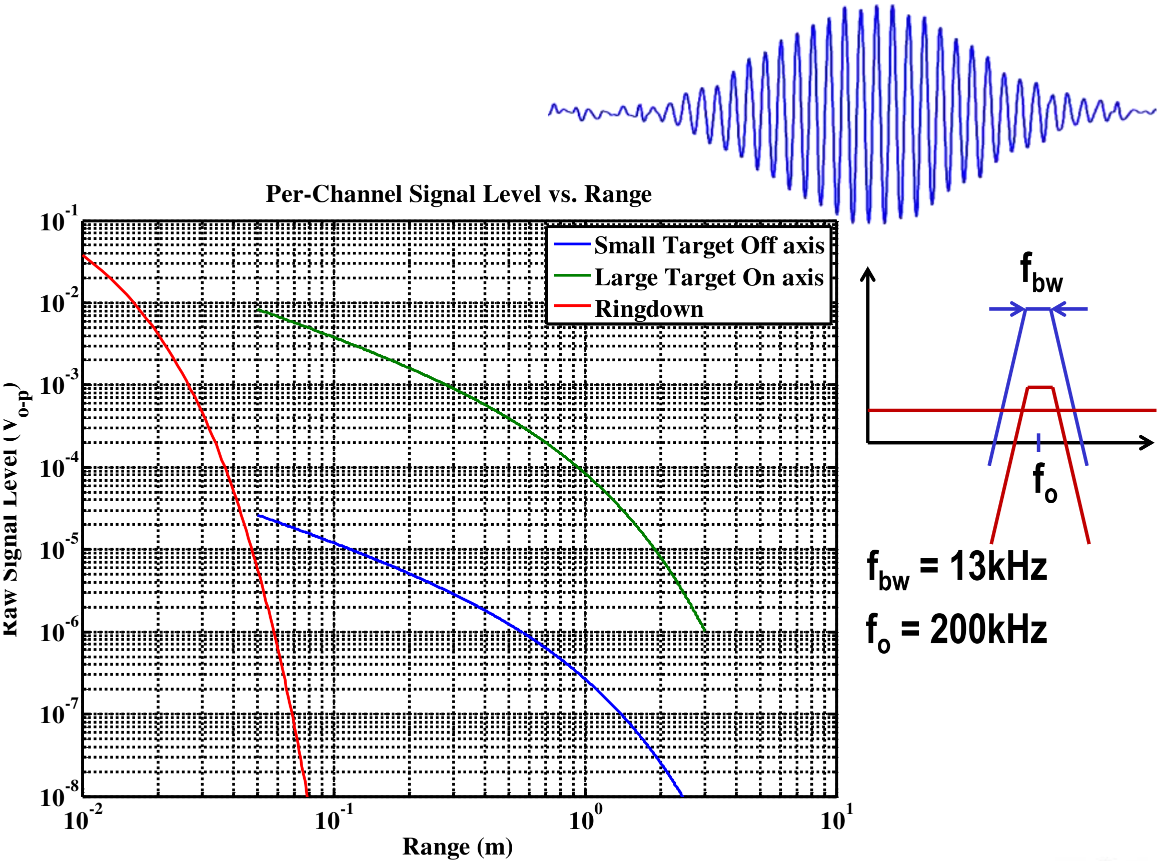


Figure 5.2: Left: Received input signal vs. range for a typical transducer. Right: Signal input spectrum, blue, and shape of noise spectrum, red. Top: typical shape of received voltage pulse.

Another important consideration for the front–end is the input capacitance of the front–end. Whether or not the front–end is used in capacitive feedback, the input–capacitance of the front–end adds directly to C_o, and must be included in the above calculations. To maximize the power efficiency, we should reduce V^* until the front–end input capacitance $C_{in} = C_o$. However, in many cases this is area prohibitive, and therefore V^* should be chosen as low as possible while meeting other requirements such as offset.

Low–frequency flicker noise is also a consideration in the front–end design; for a 200kHz signal careful device sizing usually permits the $1/f$ corner frequency to be placed below the signal band.

Finally, since the SNR varies widely depending on the target location and size, the front–end can be used as a variable–gain amplifier to reduce the required dynamic range in the ADC.

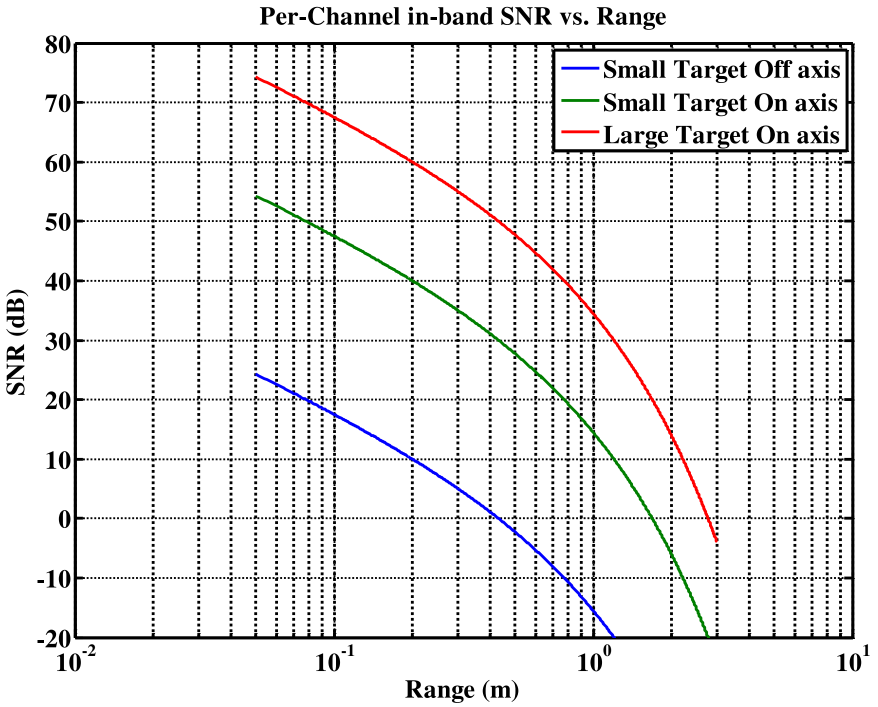


Figure 5.3: SNR versus range to target.

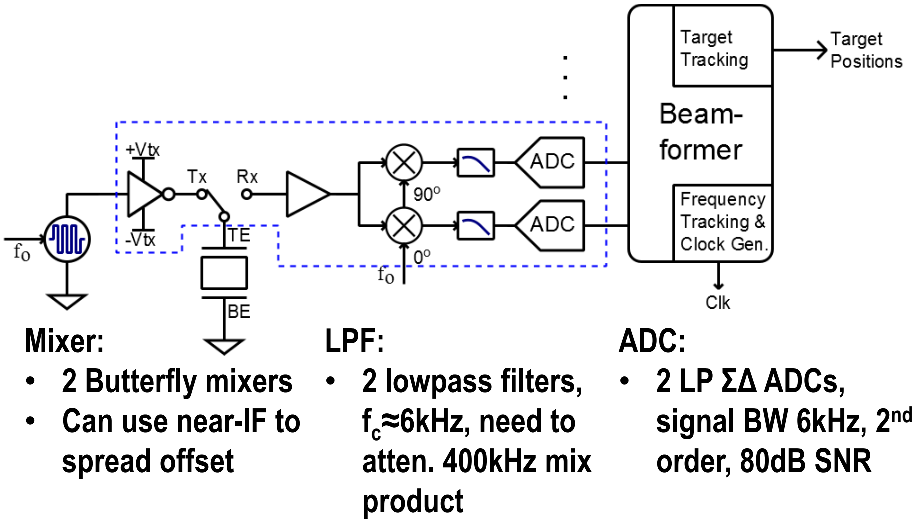
Analog to Digital Converter

The analog to digital converter (ADC) takes the signal from the front–end and converts it to a digital form that can be processed by the digital beamformer.

The most important feature of the ADC is that it should not sample the signal from the front–end amplifier until this signal is free from the wideband thermal noise generated by the amplifier. Since the front–end is designed to have a relatively high g_m in order to keep the thermal noise from this stage low, it is generally wideband. Sampling at a rate f_s will fold components above $f_s/2$ into the signal band, which can lead to dramatic noise folding of 10dB or more. An anti–aliasing filter must be incorporated before the ADC, or incorporated into the ADC itself.

A key question in the ADC architecture selection is what frequency to perform the digitization at. The signal is a relatively narrowband $\simeq$12kHz signal centered at several hundred kilohertz, so it could be digitized directly in any relatively modern CMOS process. However, butterfly mixers allow the signal to easily be translated from one frequency to another, so the designer has to decide which frequency is the best choice for digitization.

The choices can be distilled to three cases. A direct conversion architecture, such as the one illustrated in Figure 5.4, quadrature demodulates the signal to dc



Mixer:
- **2 Butterfly mixers**
- **Can use near-IF to spread offset**

LPF:
- **2 lowpass filters, $f_c \approx$6kHz, need to atten. 400kHz mix product**

ADC:
- **2 LP ΣΔ ADCs, signal BW 6kHz, 2nd order, 80dB SNR**

Figure 5.4: Direct conversion architecture, in which the signal is quadrature demodulated to dc and then converted by a low–speed ADC.

and processes the envelope of the in–phase and quadrature signals using low–speed ADCs which each need only operate with a 6kHz bandwidth. An example direct digitization architecture is shown in Figure 5.5. The ADC does not demodulate the signal, but rather digitizes it directly at the center frequency of the signal. The final case, intermediate frequency (IF), modulates the signal to a second frequency where it is digitized directly, and has a block diagram very similar to that in Figure 5.4.

There are several tradeoffs in each architecture [13], [14]. The direct conversion approach is desirable because it relies on low–speed ADCs with relatively modest requirements. However, it suffers from several problems. The OTAs in the ADC would be operating far below their $1/f$ corner frequency, so the flicker noise may degrade the SNR of the system. Worse, the demodulation signal for the butterfly mixers is likely to leak into the single–ended front–end amplifier circuit. When demodulated, this leakage will cause an unknown dc offset to appear at the input of the ADC. Since it is hard to model this leakage signal, it is difficult to design a circuit that rejects it.

The direct digitization approach does not have these limitations. The main drawback of the direct digitization approach is that the ADC has to operate at a higher frequency, thus increasing the power consumption. However, this penalty can be mitigated with careful choice of ADC architecture.

The successive–approximation (SAR) and the ΣΔ architectures are commonly used in this frequency range. The SAR architecture is a Nyquist rate converter that is clocked at $N_{bits}f_s$ in order to resolve a signal with N_{bits} of resolution. For $f_s \simeq$ 500kHz, a 12 bit converter would run at 6MHz.

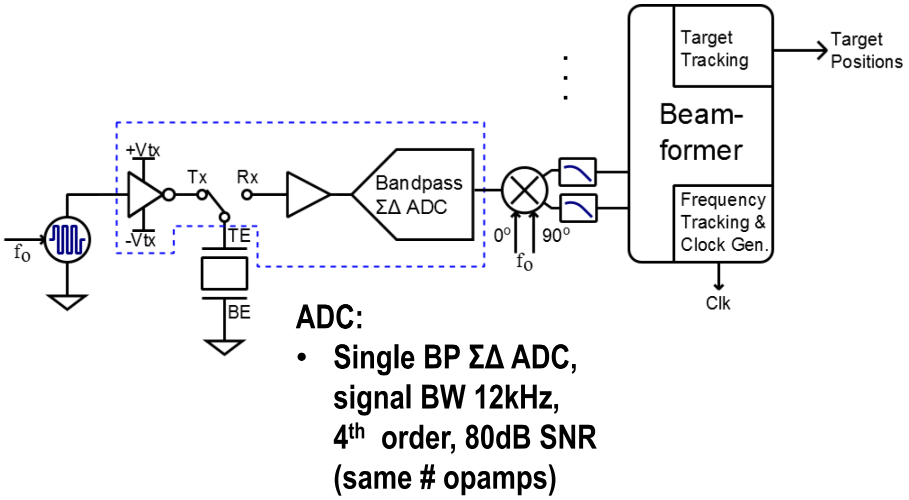


Figure 5.5: Direct digitization architecture, in which the signal directly converted by an ADC working at the center frequency f_o.

A ΣΔ converter [15] derives its resolution from oversampling. A one-bit comparator converts the output of a loop filter into 1-bit data, resulting in high quantization error. However, this 1-bit data is fed back and subtracted from the input signal, and then passed through the loop filter. For a 2nd order converter with integrators in the loop filter, the noise zeroes are placed at dc, and the required oversampling ratio is about 120, which leads to a sampling rate of 24MHz.

The bandpass ΣΔ ADC architecture overcomes these limitations by placing resonators in the forward path with their resonance centered at the signal frequency. The high in–band gain means that the quantization noise from the comparator is pushed away from the signal band. In this way, the f_s of the converter can be reduced dramatically.

The resonators in the loop–filter should be matched to the transducer's resonant frequency, which varies due to fabrication tolerances. Switched–capacitor filters lend themselves to this tuning capability, since the resonant frequency of a switched–capacitor resonator is set by the sampling frequency f_s and ratios of capacitors which can be well–defined in typical CMOS processes.

For a fourth–order bandpass ΣΔ with a single–bit quantizer and two resonators at f_o, the signal to quantization noise ratio (SQNR) is [15]:

$$SQNR \simeq \frac{15M^2\mathrm{OSR}^2}{2\pi^4}, \tag{5.3}$$

where M is the full–scale amplitude and $\mathrm{OSR} = \frac{f_s Q}{f_o 2}$ is the oversampling ratio. In our implementation, $f_s = 16 f_o \simeq 3.2$MHz, yielding $\mathrm{OSR} = 120$, was chosen to provide good enough SQNR without making the required capacitor ratios too large.

Energy Tradeoffs

For the purposes of energy tradeoffs, we adopt a conservative energy estimate of 2pJ/conv–step for the ADC. The SNR is a function of the power burned in the front–end as discussed in the previous sections. As the SNR increases, more resolution is required in the ADC, so the ADC power increases as well. The transmit energy is defined in the previous sections. The energy per measurement, E_{meas}, is the transmit energy plus the power burned in the receiver multiplied by the receive time, which depends on the maximum range and therefore the SNR. So we have:

$$E_{meas} \simeq QC_oV_{TX}^2 + \frac{2R_{max}}{c}\left(E_{cs}\sqrt{\mathrm{SNR}_{\mathrm{max}}} + P_{fe}\right), \tag{5.4}$$

$$\mathrm{SNR}_{\mathrm{max}} = \frac{V_{in,max}}{\frac{k_bT}{L_M}\frac{1}{\omega_o^2C_o^2}\left(1 + \frac{R_M\omega_o^2C_o^2}{2P_{fe}}V_{DD}V^*\right)}. \tag{5.5}$$

Figure 5.6 shows the energy per measurement for 3 different levels of transmit voltage. The asymptotic behavior at low maximum ranges is caused by the energy consumption being dominated by the transmit energy. The Brownian motion of the air and the transmit pressure set an absolute maximum achievable range, and this causes the asymptotic behavior at the top of the graph, where the transducer noise is dominant and thus increasing the power consumption of the front–end does not increase the SNR further.

As you can see, increasing the transmitter voltage increases the maximum possible range. Because the transmit energy is only a small part of the total, for medium range designs it may be desirable to increase the transmit voltage and reduce the sensitivity of the receiver. On the other hand, 5V options are more common typical CMOS processes than 12V or 30V options, and this plot shows that 1m range is achievable with only 5V transistors.

For the implementation presented here, we chose to maximize performance within reason to demonstrate what the technology is capable of. For this reason, the target design is on the red curve, at approximately 1.5m maximum range and 2uJ/ch energy per measurement.

5.2 Receive circuit design

This section briefly highlights the design of the receiver circuit. The basic block diagram is shown in Figure 5.7.

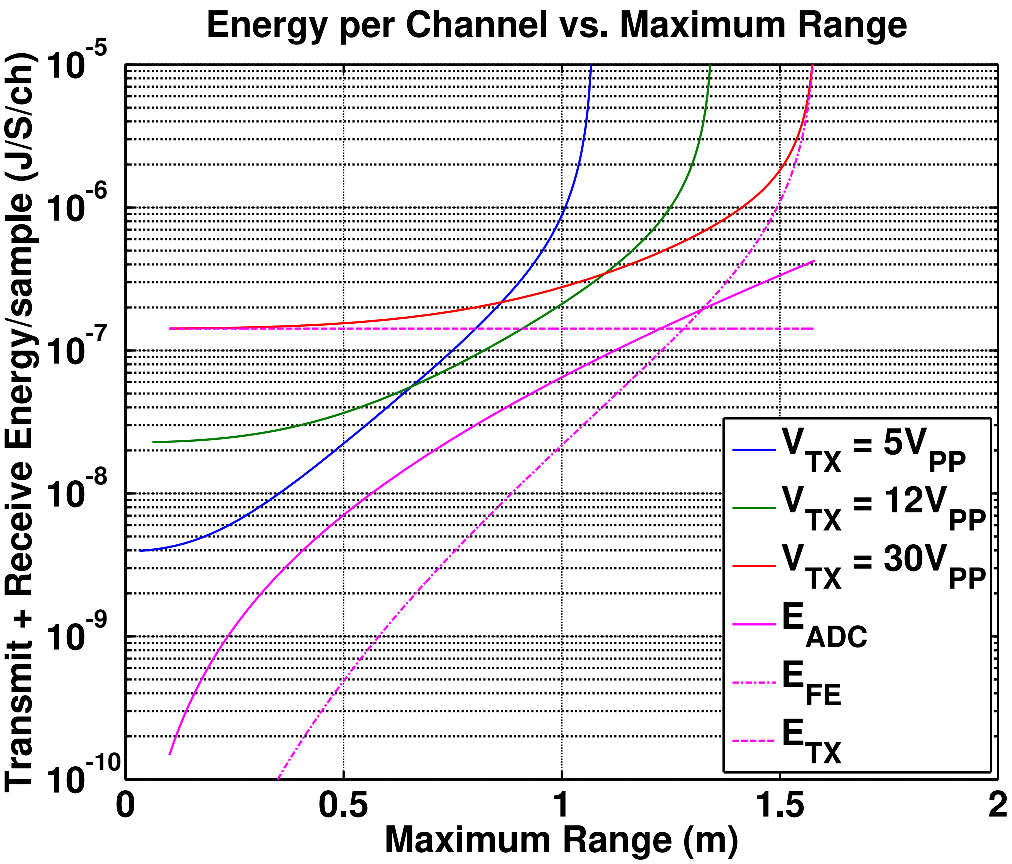


Figure 5.6: Energy consumption per measurement versus maximum desired range. Increasing maximum range requires increased sensitivity in the receiver plus longer measurement time, leading to dramatically increased energy consumption.

Front–end

The front–end amplifier was designed to be power–efficient and to provide a wide variable gain tuning capability. The amplifier is a differential OTA which uses both NMOS and PMOS input pairs to increase the current efficiency. Each unit cell was designed to have a g_m of 12μS, and 64 unit cells are used, so the g_m is tunable between $12\mu\text{S} - 768\mu\text{S}$. The V^* of the input transistors was chosen to be 0.1V, which yields a current of 600nA per unit cell.

Figure 5.8 shows the unit cell used in this work. The use of NMOS and PMOS input pairs, sized for equal $g_m = 6\mu$S, doubles the current efficiency of the design.

The front–end layout is common–centroid, with dummy cells around the periphery.

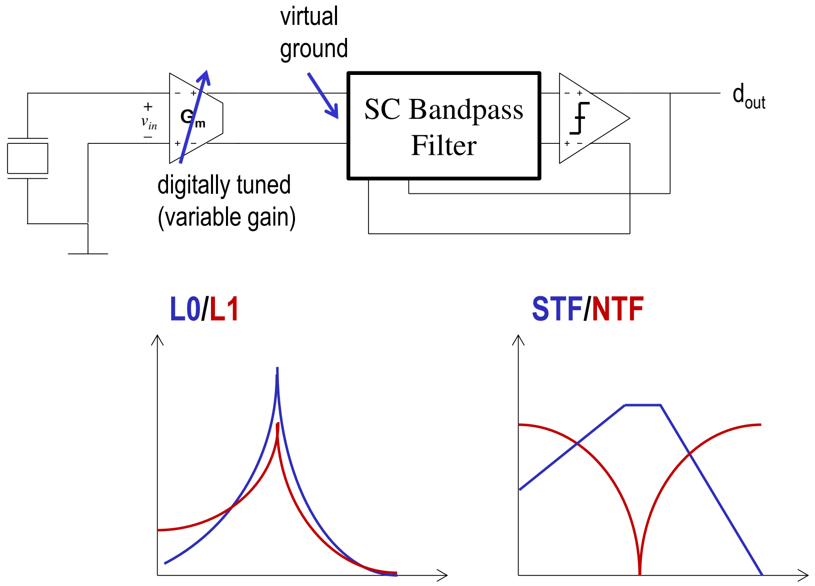


Figure 5.7: Schematic view of the receiver showing variable gain front–end and bandpass ΣΔ ADC.

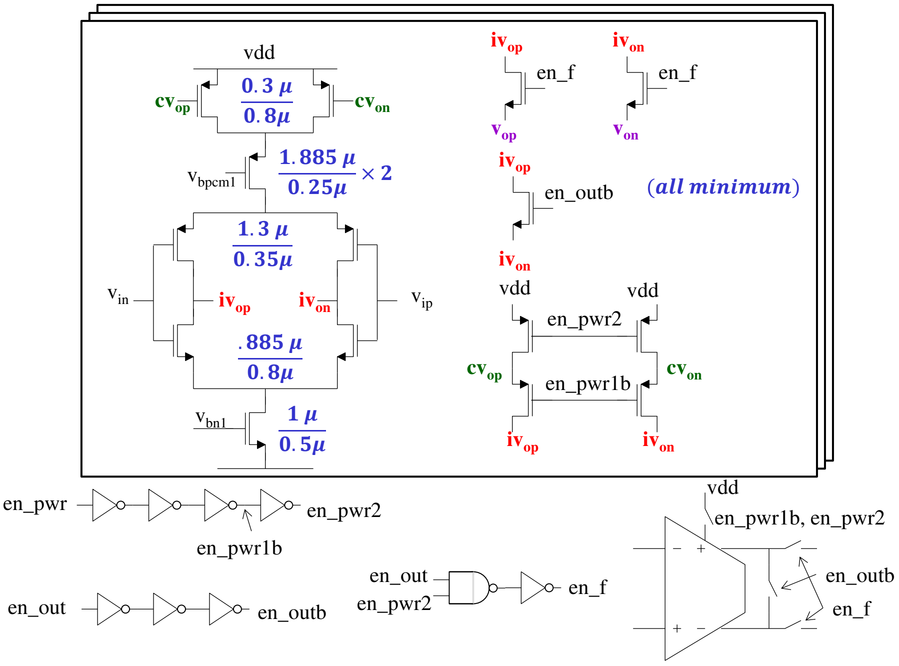


Figure 5.8: Unit cell OTA used in front–end amplifier.

Bandpass ΣΔ ADC

The bandpass ΣΔ ADC was designed with the help of the DSMtoolbox MATLAB functions designed by Richard Schreier [15]. A cascaded–resonator feedback (CRFB) topology was chosen because the capacitor ratios in this topology are not very high.

An important improvement compared to the conventional CRFB topology was the choice not to use an input sampling capacitor. A sampling capacitor at the input would alias wideband noise from the front–end into the band of interest. As mentioned in the previous section, it is easy to incur a 10dB penalty through noise folding. Instead of a sampling capacitor at the input, the input current from the front–end amplifier is integrated in the feedback capacitor of the 2nd integrator of the 1st resonator. In this manner, the signal is not sampled until after it has been filtered by the 2nd integrator, thereby preventing noise folding. The 2nd integrator thus functions as a continuous–time amplifier for the purposes of the signal, while processing the feedback signal from the comparator and the input signal from the 1st integrator in a discrete time fashion.

Figure 5.9 shows a simplified schematic of the front–end and the first resonator. For clarity the continuous–time implementation is shown. A more detailed schematic showing the switches and phasing is presented in Chapter 6.

The resonant frequency of the resonator shown in Figure 5.9 is controlled by the time constants of the two integrators:

$$f_{res} \simeq \frac{1}{(2\pi)^2} \frac{1}{R_{r1} C_{i1} R_{s2} C_{i2}} = \frac{f_s^2}{(2\pi)^2} \frac{C_{r1} C_{s2}}{C_{i1} C_{i2}}. \tag{5.6}$$

This equation explains the usefulness of the switched–capacitor amplifier in one line. The resonant frequency depends on f_s, which can be locked to an accurate clock, and the ratios of capacitors, which can be very accurately controlled in CMOS technology.

The DSMtoolbox permits the designer to calculate capacitor ratios in a straightforward manner. We conservatively used 5 of the smallest MIM capacitors available in our process to make the unit capacitance. The MIM capacitor layout is common–centroid with dummy capacitors around the periphery for best matching.

A key drawback to this architecture is susceptibility to dc offset. The signal resides at 200kHz, so we want to reject any dc signal, which the architecture does since the resonator rejects dc signals. However, if the front–end has an offset V_{off}, this translates into a worst–case input offset current $i_{off} = V_{off} g_{m,max}$ which is integrated onto C_{i2} and then canceled by the first integrator. As a result, the output of the first integrator must contain a dc voltage V_1 which produces a dc current through C_{s2} to cancel i_{off}. This dc voltage acts to reduce the allowed swing of the integrator before it saturates. If we choose a maximum V_1 that we want to sacrifice to offset cancellation, then we have

$$C_{s2} = \frac{V_{off,max}}{V_{1,max}} \frac{f_s}{g_{m,fe}}. \tag{5.7}$$

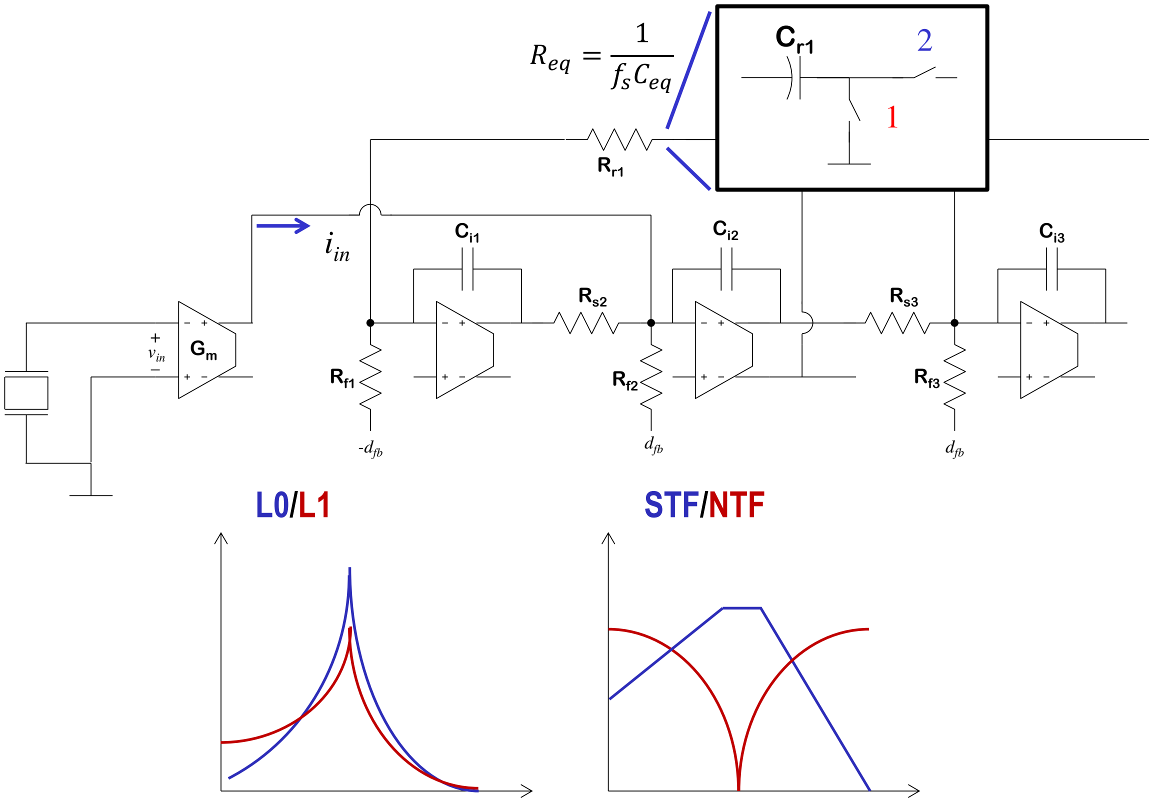


Figure 5.9: Schematic of front–end interface to 1st resonator in ΣΔ ADC. For clarity, the switched capacitor resistors in the design have been replaced with resistors, and the fully–differential structure is shown single–ended. The plots at bottom show (left) L0, the gain from the input current to the output of the loop filter, and L1, the gain from the one–bit feedback to the output (right) the signal transfer function and the noise transfer function.

As a result, C_{s2} must be sized relatively large to reject dc offset due to random mismatch in the front–end.

The OTAs in the ADC are sized for 9τ settling within half of the sampling period. This puts a requirement on the g_m of the OTAs which sets the power consumption. Because of the dc offset issue, the capacitors in the first resonator are relatively large, and the OTAs in the first resonator had to be scaled up by a factor of 3 as a result.

We built a simulink model to simulate the converter using MATLAB. We used this model to simulate the ac response and time domain behavior of the converter, determine the signal levels at each integrator, and fine tune the capacitor ratios. Figure 5.10 shows the simulink model. After the capacitor ratios are determined, the OTAs are sized, and the transistor–level schematic is designed and simulated using

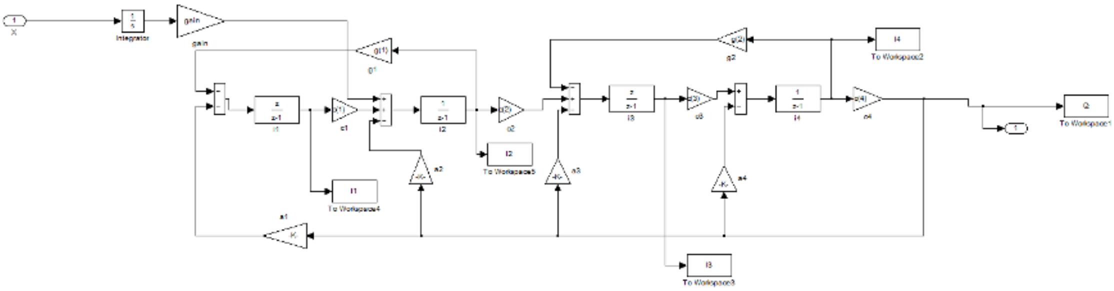


Figure 5.10: Simulink model for front–end amplifier and ADC. The 2nd integrator is split into a continuous time integrator for the input signal and a discrete time integrator for the feedback signals.

Spectre. Finally, we lay out the design and extract it and simulate the extracted layout again using Spectre. Two example spectrums are shown in Figure 5.11 from the beginning and end of the process.

This chapter presented the design procedure for a readout integrated circuit for a 3D rangefinder. Results and implementation details are presented in the next chapter.

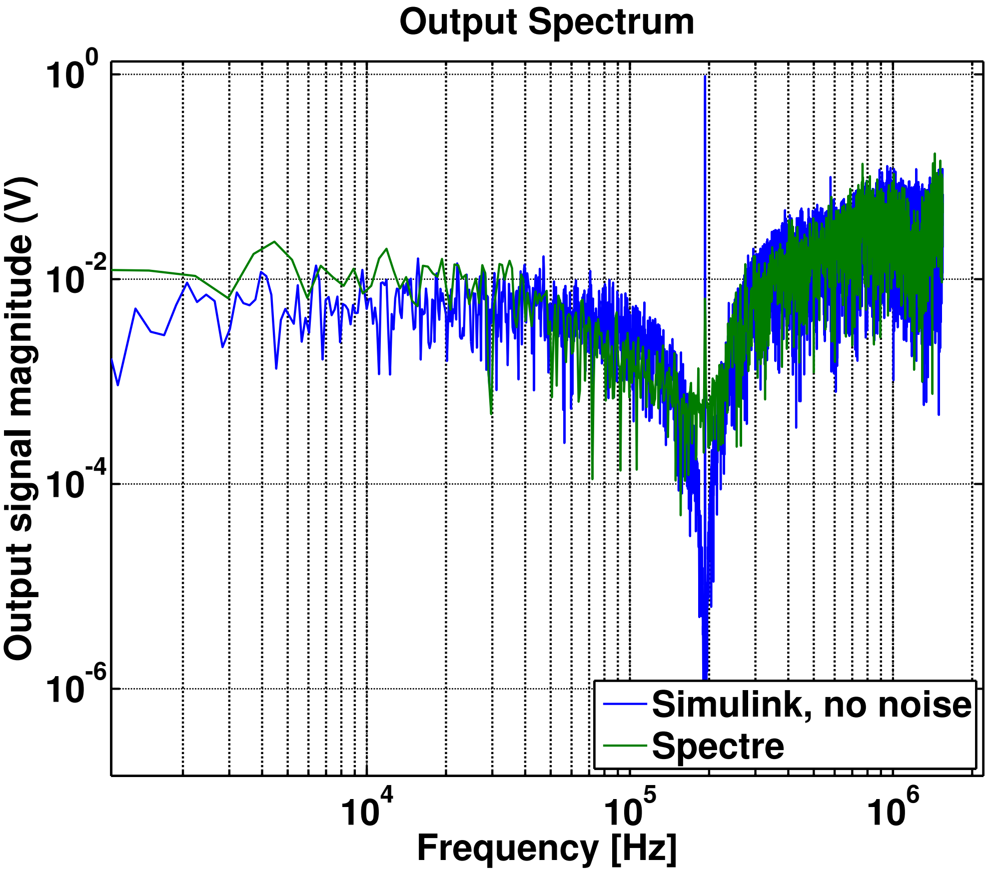


Figure 5.11: Output spectrum of the ΣΔ ADC with a full–scale signal, simulated in simulink, and the minimum–detectable signal, simulated with noise post–layout.

Chapter 6

Implementation of a 3D Ultrasonic Gesture Recognition System on a Chip

This chapter presents an ultrasonic 3D gesture recognition system which uses a custom transducer chip and ASIC to sense the location of targets such as hands. The system block diagram is shown in Figure 6.1. Targets are localized using pulse-echo time-of-flight methods. Each of the ten transceiver channels interfaces with a MEMS transducer, and each includes a transmitter and a readout circuit. Echoes from off-axis targets arrive with different phase shifts for each element in the array. The off-chip digital beamformer realigns the signal phase to maximize the SNR and determine target location.

6.1 Ultrasound Chip

The 450μm diameter piezoelectric micromachined ultrasound transducers (pMUTs) used in this work are made up of a 2.2μm thick AlN/Mo/AlN/Al stack deposited on a Si wafer and released with a back-side through-wafer etch. The bottom electrode is continuous, while each pMUT has a top electrode lithographically defined to actuate the trampoline mode. Each pMUT can transmit and receive sound waves, and is operated at its resonance of 217kHz ± 2kHz with a bandwidth of 12kHz. The impedance of the transducers is dominated by 10pF transducer capacitance, and the motional resistance at resonance is 2.4MΩ. The resonant frequencies of the pMUTs vary due to fabrication, temperature, and packaging stress, so online frequency tracking is used to maintain maximum signal to noise ratio during operation.

Two pMUTs are used for transmission and seven for reception in the array shown in Fig. 6.2. The receive array is 3.5 wavelengths wide in the x-angle axis, allowing

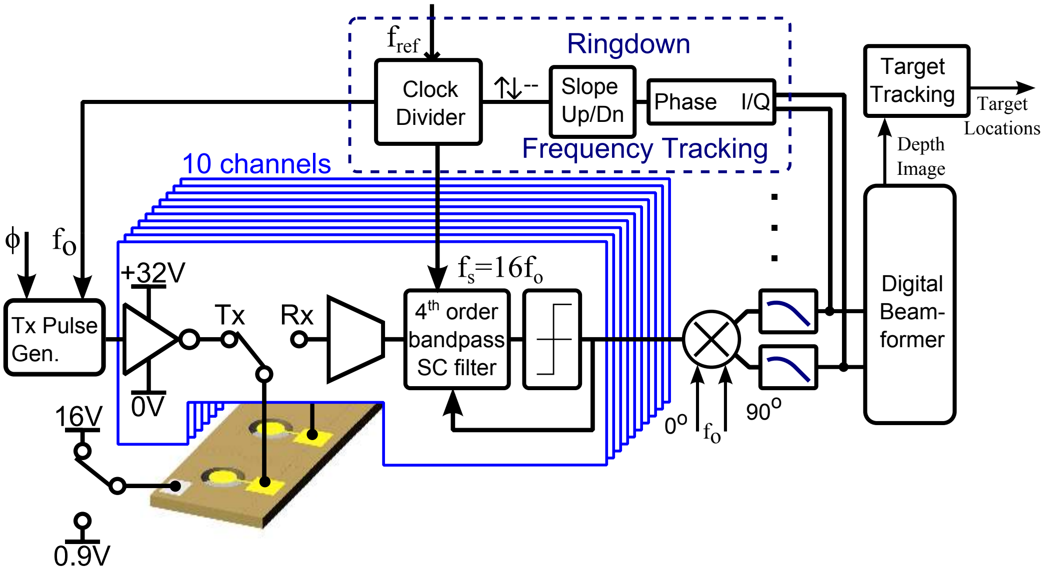


Figure 6.1: Block diagram of ultrasonic gesture recognition system.

targets separated by more than 15^o to be distinguished. In the y-angle axis the array is only 0.16 wavelengths wide, sufficient to determine the y-angle to the target by measuring the average phase difference along the y axis of the array. The center element of the receive array and the element 900μm above it are used to launch a 138μsec24mm long pulse of sound into the environment. The transmit configuration illuminates a wide field of view, permitting the capture of an entire scene in a single measurement. Applications requiring better target resolution or greater maximum range can also use transmit beamforming at the expense of reduced measurement rate.

6.2 Interface Electronics

Each cycle begins with the launch of an acoustic pulse. Figure 6.5 shows the schematic of a single channel. High voltage level shifters actuate the S_{TX} transmit switches, setting the transducers bottom electrode to 16V to permit bipolar actuation of the transducer. The transmitter then excites the transducer with a $32\mathrm{V}_{\mathrm{PP}}$ square wave for 30 cycles at the transmit frequency f_{TX} which is locked to $\frac{1}{16}$ of the sampling frequency f_s.

The transducer's nonlinearity reduces the maximum output achievable with large transmit voltages. Figure 6.3 shows the measured SNR for a target at 300mm for different transmit voltages. The transducer's nonlinearity reduces the SNR by 4dB

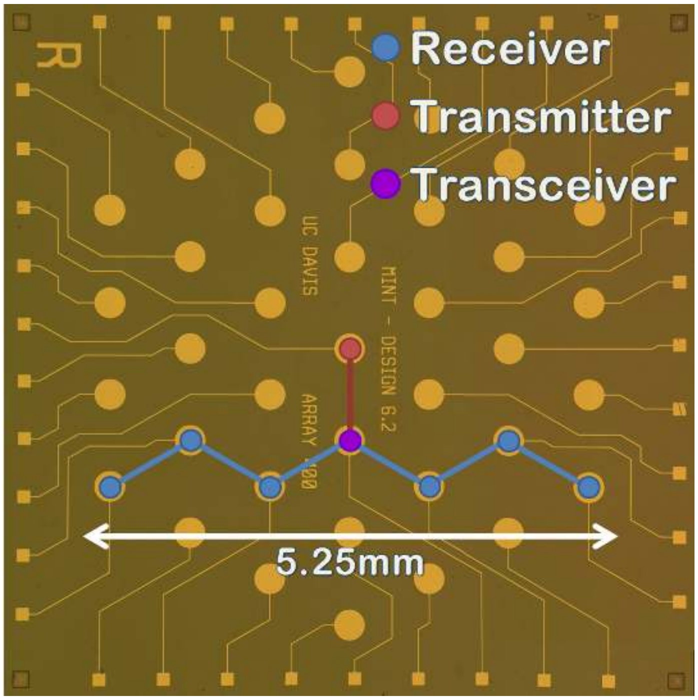


Figure 6.2: Chip micrograph of MEMS ultrasound chip showing array configuration.

when using a 32$\mathrm{V_{pp}}$ excitation.

In ultra–low power rangefinders, where the transmit energy is a large part of the energy budget, a designer may choose to use a reduced transmit voltage in order to use the transmitter in the linear regime where it is more efficient. Figure 6.4 shows the measured SNR as a function of the transmit energy. The transmitter efficiency of this implementation was measured to be 30%.

At the end of the transmit phase, the mechanical energy stored in the inertia of the pMUT dissipates and the pMUT rings down at its natural frequency. The S_{RX} receiver isolation switches are turned on, and a resistor converts the ring down current to a voltage which is subsequently amplified and digitized by the receiver normally. The ring-down signal is then I/Q demodulated with f_{TX}. The slope of the phase signal during the ring down indicates the frequency offset and is used to update the f_s and f_{TX} used in the next measurement. Figure 6.6 shows the operation of the frequency tuning loop. An initial 57kHz offset frequency is nulled to 1kHz within 30 measurement cycles.

After 86μsec, the ring down signal has decayed sufficiently for the S_{ring} switch to be opened, beginning the processing of received echoes. At this point, the signal from the

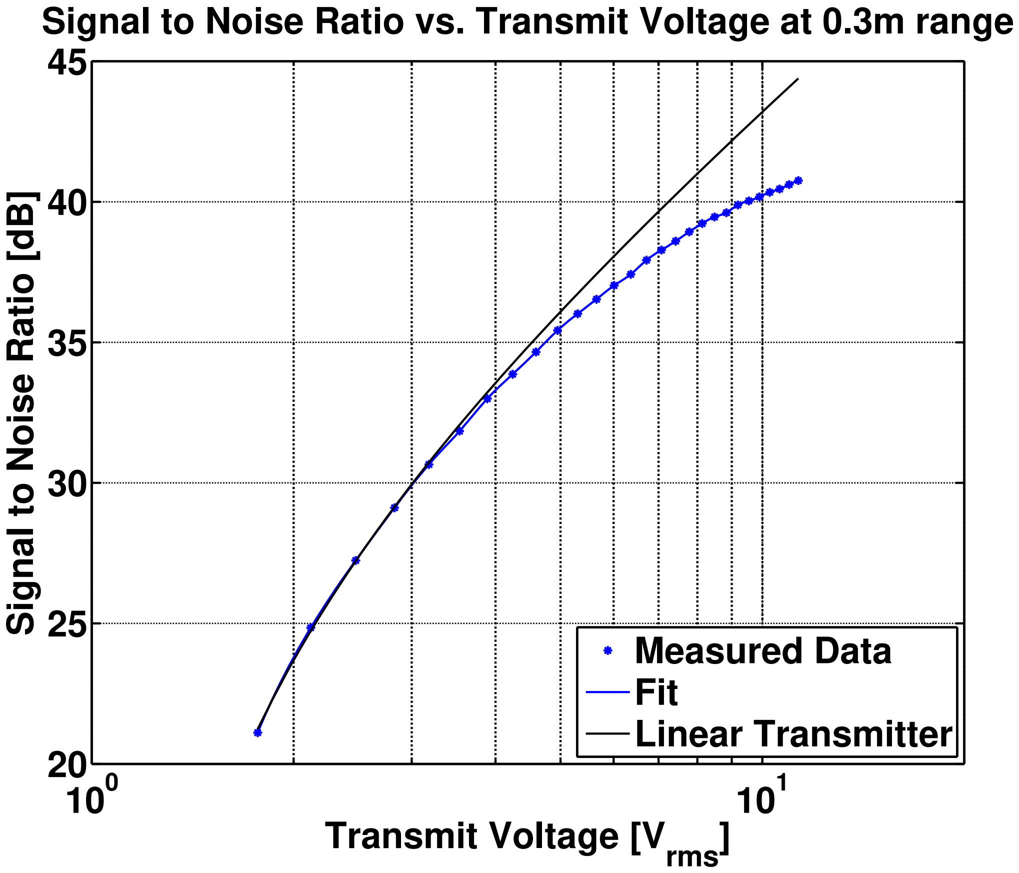


Figure 6.3: Signal to noise ratio for a target at 300mm vs. transmit voltage.

transducer is integrated on the transducers capacitance, and the front-end measures a voltage which is proportional to the displacement of the transducers membrane.

The front-end amplifier consists of an open-loop current-reuse OTA with both NMOS and PMOS differential pairs biased near subthreshold for current efficiency. The front-end current is integrated onto the integrating capacitor of the second stage, which also makes up an integrator in the first of two switched capacitor resonators. Although the second stage is a switched capacitor integrator, the front-end current is processed in a continuous time fashion before it is sampled at the output of the second integrator. As a result, the second integrator acts as an anti-aliasing filter for the wideband noise generated by the front-end and prevents this noise, the dominant noise source in the receiver architecture, from being aliased into the band of interest.

The signal then passes through a second switched capacitor resonator and is quantized by a comparator. The high in-band gain provided by the 4th order bandpass filter shapes the wideband quantization noise to be away from the signal at f_{TX}. The SC resonators are designed to resonate at $\frac{1}{16}$ of the sampling frequency, and which

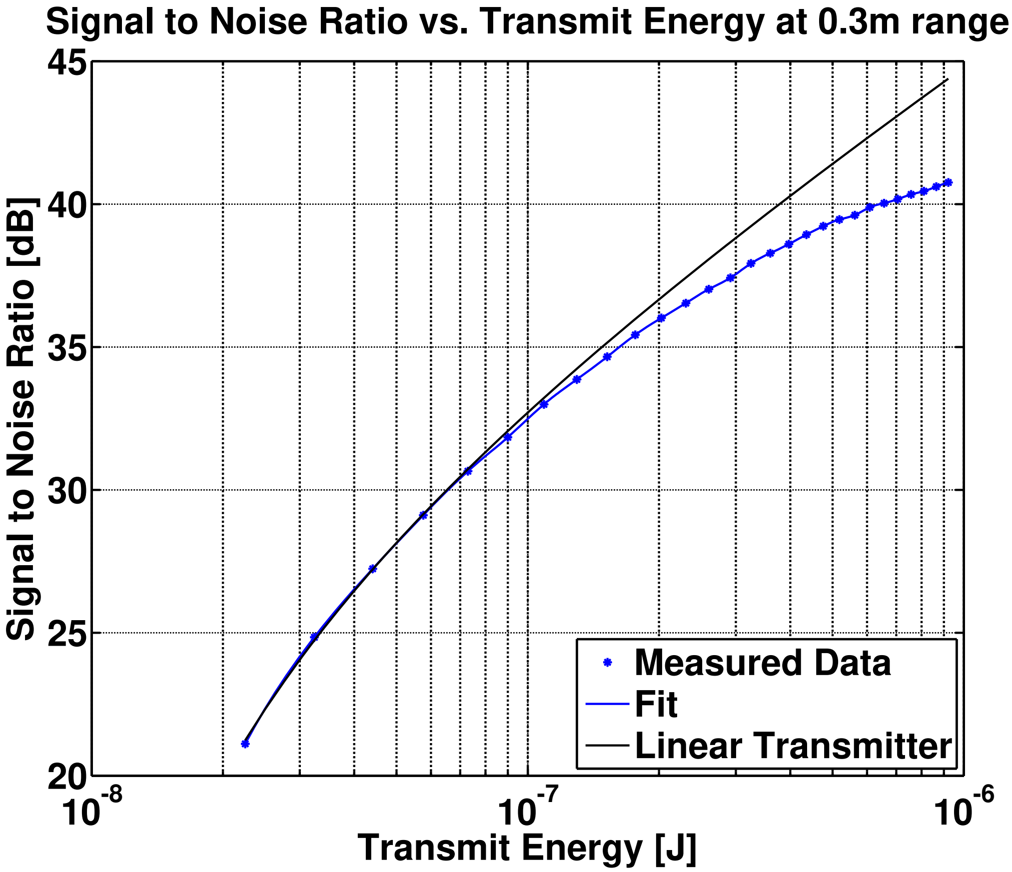


Figure 6.4: Signal to noise ratio for a target at 300mm vs. transmit energy.

is tuned to the resonance of the transducer using the ring down measurement by adjusting f_s.

The output of each ΣΔADC is I/Q demodulated, filtered, and downsampled off-chip. A digital beamformer [11] processes the signals to maximize the receive SNR and determine the x-angle location of the target. This process can be repeated in the orthogonal angle axis to implement 3D beamforming; in this work we forgo 3D beamforming since the tiny y-axis aperture does not provide any y-axis resolution.

Thermal noise in the front-end amplifier and the thermal motion of air limit the minimum detectable echo. The input referred noise of the amplifier is $11\text{nV}/\sqrt{\text{Hz}}$, and the noise voltage of the transducer is $6\text{nV}/\sqrt{\text{Hz}}$ at resonance. Fig. 6.7 shows the measured signal–to–noise ratio vs. range for a 127mm x 181mm flat rectangular target.

The measured data from the SNR plot in Figure 6.7 can be combined with the measured SNR versus transmit voltage data in Figure 6.3 to extrapolate the maximum

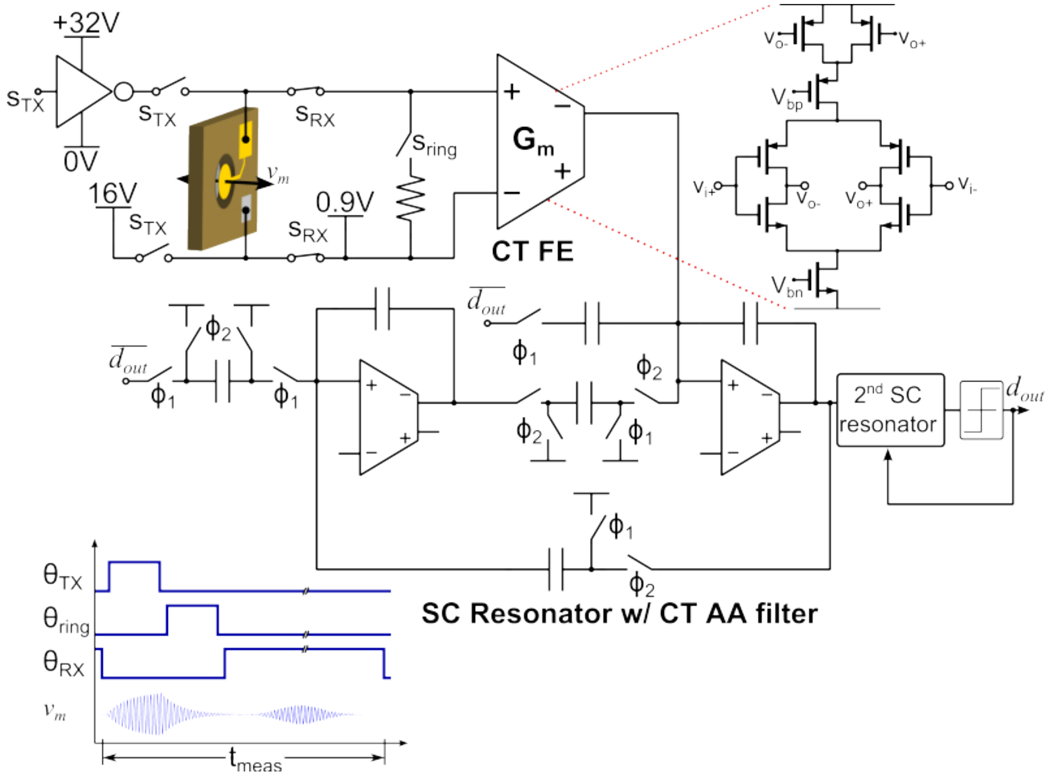


Figure 6.5: Schematic of readout circuit.

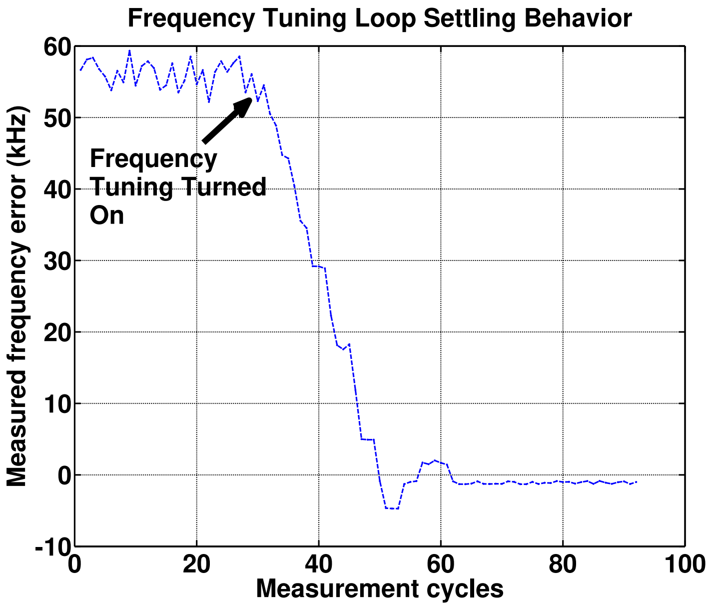


Figure 6.6: Settling behavior of frequency tuning loop.

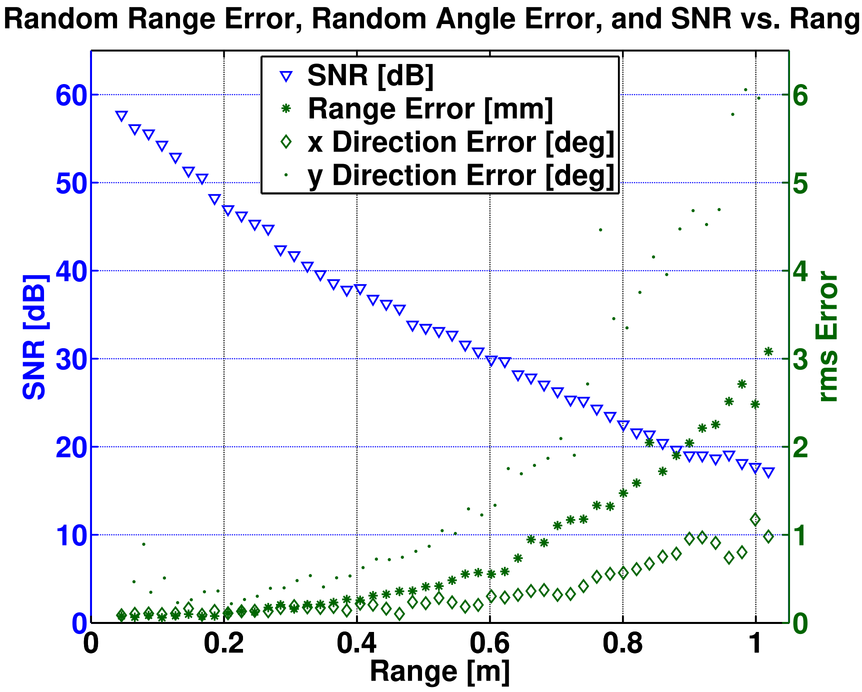


Figure 6.7: Signal to noise ratio and target localization accuracy vs. range for 127mm x 181mm flat rectangular target.

range for different transmit voltages. Figure 6.8 shows the extrapolated maximum range versus transmit voltage, with 12dB SNR set as the maximum range. As you can see, reducing the transmit voltage by half only reduces the maximum range about 15%. Systems that are limited to lower voltage transistors or wish to operate with higher transmit efficiency will choose to operate with lower transmit voltage than is used in this implementation.

Figure 6.7 also shows the rms error in the range and direction measurement. Amplitude noise in the received signal limits the accuracy of the time-of-flight estimate. Fig. 6.9 shows the echo captured from a users hands and head as he poses as shown. The system tracks objects between 45mm to 1m away and over an angular range of ±45°. Echoes from targets at a range of 1m return after 5.8ms, and this sets the maximum measurement rate of the system.

Figure 6.10 shows a micrograph of the readout IC which was fabricated in a 0.18μm CMOS process with 32V transistors. For a 1m maximum range, the system presented here uses 13.6μJ per measurement. At 30fps, the receive power consumption is 335μW and the transmit power consumption is 66μW. The energy consumption scales roughly linearly with maximum range. For a maximum range of 0.3m, the

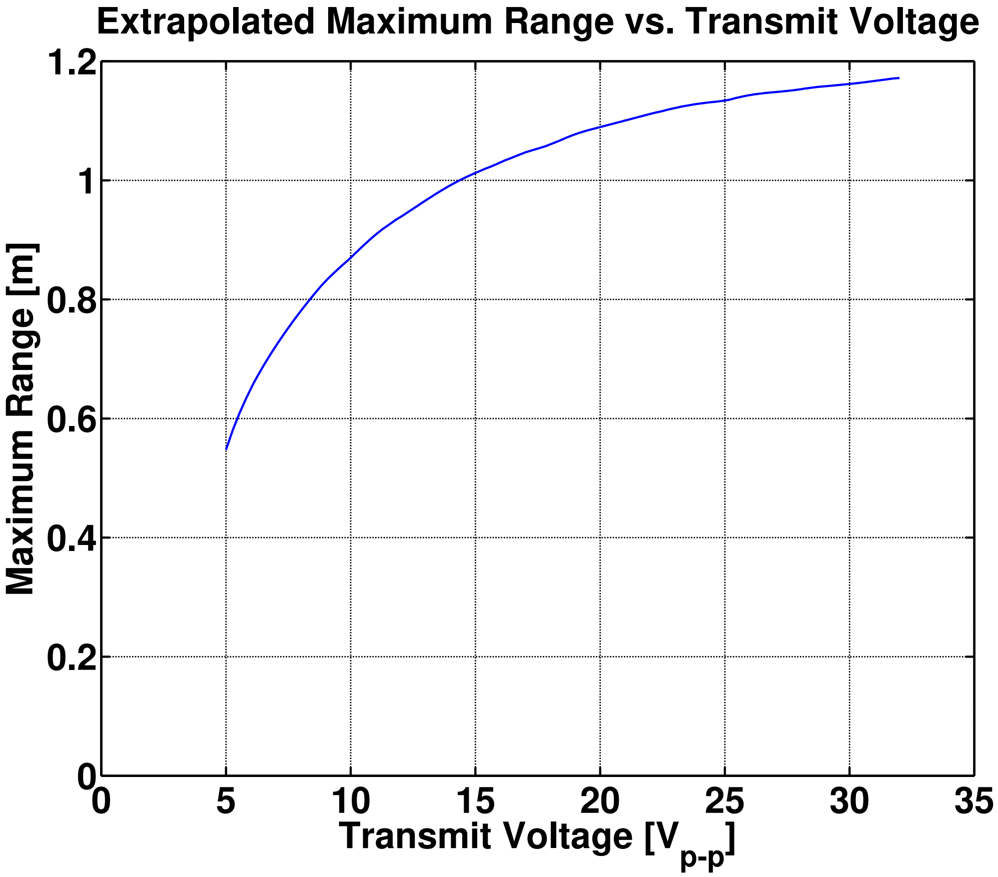


Figure 6.8: Maximum range versus transmit voltage, extrapolated from measured data in Figure 6.7 and 6.3.

energy per frame is reduced to $< 0.5\mu$J per frame. Single-element range measurements can be conducted at 10fps using only 5μW.

Table 6.1 compares the performance of this system to an earlier MEMS ultrasonic 1D rangefinder [8] and two recent optical 3D rangers [16], [17]. This ultrasonic 3D rangefinder offers dramatically reduced energy consumption compared to optical methods while permitting 3D target tracking. The energy consumption trades off with performance, permitting continuous operation in even tiny mobile devices. These characteristics enable energy-efficient gestural interfaces in applications such as smartphones and tablets, and permit gestural user interfaces in tiny mobile devices too small to accommodate a conventional touchscreen.

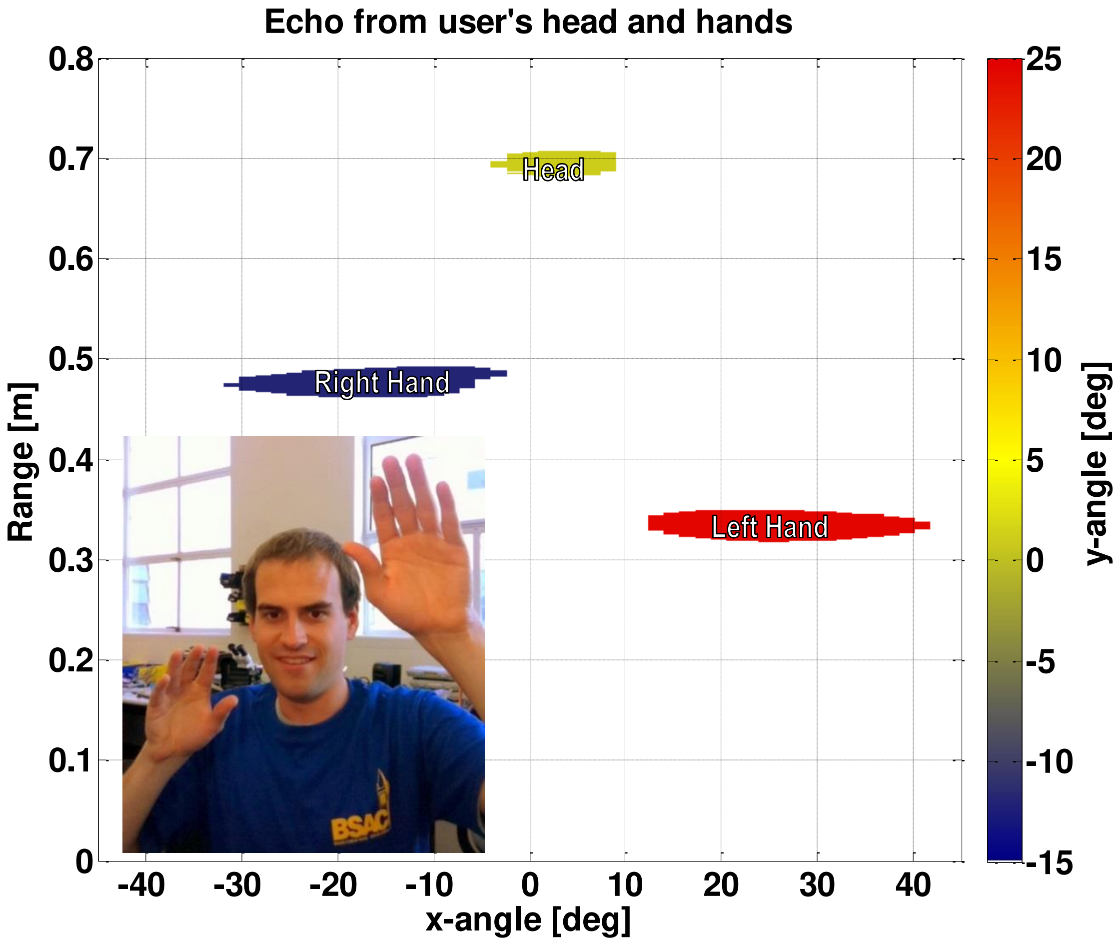


Figure 6.9: Echo from users hands and head when posing as shown. Color axis shows y-angle position of the targets. Beamformed data is thresholded at 12dB SNR.

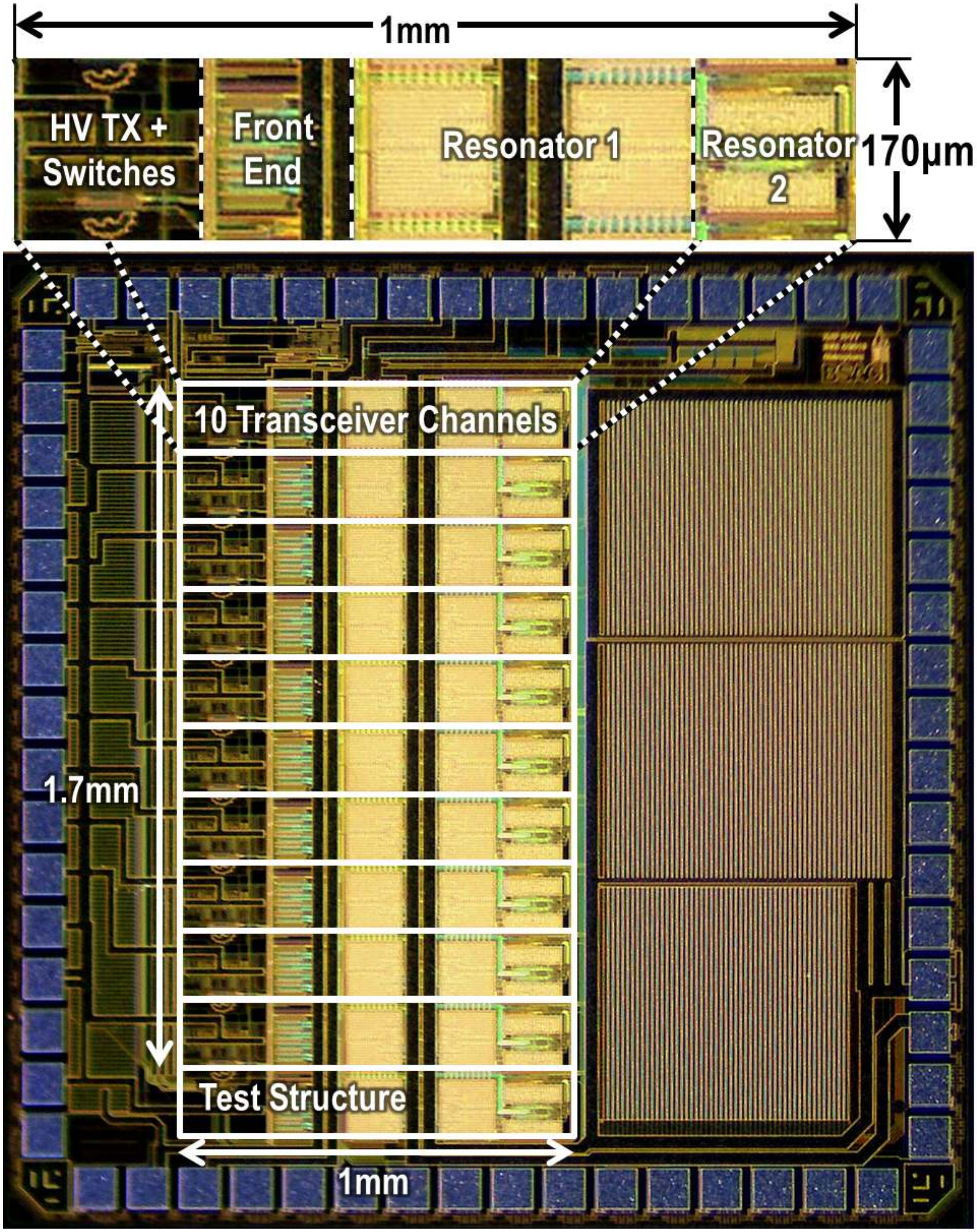


Figure 6.10: Chip micrograph of 10 channel ultrasonic gesture recognizer readout ASIC.

Table 6.1: Performance Comparison

	This work	[8]	[16]	[17]
Method	Time-of-flight sound	Phase-shift sound	Phase-shift light	Phase-shift light
Transducers	AlN pMUT array	Si Thermal Ultrasound	Pinned Photodiode	Avalanche Photodiode
Carrier wavelength / Modulation wavelength	1.6mm / 48mm	3.6mm / 250mm	850nm / 15m	850nm / 12m
CMOS Process	0.18μm	0.8μm	0.13μm	0.35μm
Min / Max Range Demonstrated	45mm / 1m	18mm / 0.11m	200mm / 7m	500m / 5m
Range error	0.41mm$_{\mathrm{rms}}$ @ 0.5m	2.5mm @ 0.1m	7mm$_{\mathrm{rms}}$ @ 1m	19mm$_{\mathrm{rms}}$ @ 2m
Field of View (x / y) (°)	90° / 90°	No angle measurement	Not specified	18° / 18°
Multi-target angular resolution (x / y) (°)	14.8° / NA	Single target only	(480px x 360px)	0.28° / 0.28° (64px x 64px)
Max Rate of Measurement	170 fps	50 fps	100 fps	200 fps
Energy per Measurement	13.6μJ	6.5mJ	20mJ	>> 2.9mJ

Chapter 7

Conclusion

Rangefinders, radios, and proximity sensors all share something in common: they transmit and receive signals from the environment, and hope to learn something from those signals–either about what was sent or about the environment itself.

Today, the overwhelming majority of the billions of microsystems that send or receive signals from the environment use electromagnetic radiation. This work shows that this need not always be the case, and that the use of sound as an alternative to light enables the creation of tiny, low–power rangefinders which operate with less than 1/100th of the power consumption of state–of–the–art optical rangefinders.

This dissertation presents an ultrasonic rangefinder which uses piezoelectric micromachined ultrasonic transducers and custom CMOS readout electronics to perform 3D rangefinding. The primary contribution of this work is to combine methods and theory from acoustics, radar, mechanics, and circuits in order to create a methodology for ultrasonic rangefinder design.

This work is easily extended to other ultrasonic microsystems that will use similar transducers. Already, there is work ongoing on ultrasonic fingerprint sensors. For air–coupled ultrasonic systems, there are several logical avenues of research. A key limitation of the current system is the angular resolution, and it is possible that new operating modes could improve the angular resolution substantially without dramatically increasing the number of sensors. There are also several applications which could be very fruitful, including indoor navigation and low–power data transmission. These applications have many parallels with 3D rangefinding, but their feasibility and utility is a matter for future research.

Bibliography

[1] R. Przybyla, S. Shelton, A. Guedes, R. Krigel, D. Horsley, and B. Boser, "In-air ultrasonic rangefinding and angle estimation using an array of aln micromachined transducers," in *Hilton Head Solid-State Sensors, Actuators and Microsystems Workshop 2012*, 2012, pp. 50–53.

[2] S. Shelton, M.-L. Chan, H. Park, D. Horsley, B. Boser, I. Izyumin, R. Przybyla, T. Frey, M. Judy, K. Nunan, F. Sammoura, and K. Yang, "CMOS-compatible AlN piezoelectric micromachined ultrasonic transducers," in *Ultrasonics Symposium (IUS), 2009 IEEE International*, 2009, pp. 402–405.

[3] R. Przybyla, S. Shelton, A. Guedes, I. Izyumin, M. Kline, D. Horsley, and B. Boser, "In-air rangefinding with an AlN piezoelectric micromachined ultrasound transducer," *IEEE Sensors Journal*, vol. 11, no. 11, pp. 2690–2697, 2011.

[4] S. Senturia, *Microsystem Design*. Springer, 2000.

[5] P. Muralt, N. Ledermann, J. Paborowski, A. Barzegar, S. Gentil, B. Belgacem, S. Petitgrand, A. Bosseboeuf, and N. Setter, "Piezoelectric micromachined ultrasonic transducers based on PZT thin films," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 52, no. 12, pp. 2276–2288, 2005.

[6] D. T. Blackstock, *Fundamentals of Physical Acoustics*. John Wiley & Sons, 2000.

[7] L. Beranek, *Acoustics*, ser. McGraw-Hill electrical and electronic engineering series. McGraw-Hill, 1954.

[8] C. Kuratli and Q. Huang, "A CMOS ultrasound range-finder microsystem," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, pp. 2005–2017, 2000.

[9] S. Roehr, P. Gulden, and M. Vossiek, "Precise distance and velocity measurement for real time locating in multipath environments using a frequency-modulated continuous-wave secondary radar approach," *IEEE Trans. Microw. Theory Tech.*, vol. 56, no. 10, pp. 2329 –2339, Oct. 2008.

[10] R. Przybyla, S. Shelton, A. Guedes, I. Izyumin, M. Kline, D. Horsley, and B. Boser, "In-air rangefinding with an aln piezoelectric micromachined ultrasound transducer," *Sensors Journal, IEEE*, vol. 11, no. 11, pp. 2690–2697, 2011.

[11] M. I. Skolnik, *Introduction to radar systems.* Boston: McGraw Hill, 2001.

[12] R. Przybyla, A. Flynn, V. Jain, S. Shelton, A. Guedes, I. Izyumin, D. Horsley, and B. Boser, "A micromechanical ultrasonic distance sensor with greater than 1 meter range," in *Solid-State Sensors, Actuators and Microsystems Conference (TRANSDUCERS), 2011 16th International*, 2011, pp. 2070–2073.

[13] B. Razavi, "Design considerations for direct-conversion receivers," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 44, no. 6, pp. 428–435, 1997.

[14] A. Abidi, "Direct-conversion radio transceivers for digital communications," *Solid-State Circuits, IEEE Journal of*, vol. 30, no. 12, pp. 1399–1410, 1995.

[15] G. Schreier, R.; Temes, *Understanding Delta-Sigma Data Converters.* Wiley-IEEE Press, 2005.

[16] W. Kim, W. Yibing, I. Ovsiannikov, S. Lee, Y. Park, C. Chung, and E. Fossum, "A 1.5Mpixel RGBZ CMOS image sensor for simultaneous color and range image capture," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, 2012, pp. 392–394.

[17] O. Shcherbakova, L. Pancheri, G.-F. Dalla Betta, N. Massari, and D. Stoppa, "3D camera based on linear-mode gain-modulated avalanche photodiodes," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International*, 2013, pp. 490–491.