ABSTRACT

Title of Dissertation:	BIO-INSPIRED SONAR IN COMPLEX ENVIRONMENTS: ATTENTIVE TRACKING AND VIEW RECOGNITION
	Jacob Isbell, Doctor of Philosophy, 2021
Dissertation directed by:	Professor Timothy K Horiuchi, Department of Electrical and Computer Engineering

Bats are known for their unique ability to sense the world through echolocation. This allows them to perceive the world in a way that few animals do, but not without some difficulties. This dissertation explores two such tasks using a bio-inspired sonar system: tracking a target object in cluttered environments, and echo view recognition. The use of echolocation for navigating in dense, cluttered environments can be a challenge due to the need for rapid sampling of nearby objects in the face of delayed echoes from distant objects. If long-delay echoes from a distant object are received after the next pulse is sent out, these "aliased" echoes appear as close-range phantom objects. This dissertation presents three reactive strategies for a high pulse-rate sonar system to combat aliased echoes: (1) changing the interpulse interval to move the aliased echoes away in time from the tracked target, (2) changing positions to create a geometry without aliasing, and (3) a phase-based, transmission beam-shaping strategy to illuminate the target and not the aliasing object. While this

task relates to immediate sensing needs and lower level motor loops, view recognition is involved in higher level navigation and planning. Neurons in the mammalian brain (specifically in the hippocampus formation) named “place cells” are thought to reflect this recognition of place and are involved in implementing a spatial map that can be used for path planning and memory recall. We propose hypothetical “echo view cells” that could contribute (along with odometry) to the creation of place cell representations actually observed in bats. We strive to recognize views over extended regions that are many body lengths in size, reducing the number of places to be remembered for a map. We have successfully demonstrated some of this spatial invariance by training feed-forward neural networks (traditional neural networks and spiking neural networks) to recognize 66 distinct places in a laboratory environment over a limited range of translations and rotations. We further show how the echo view cells respond in between known places and how the population of cell outputs can be combined over time for continuity.

BIO-INSPIRED SONAR IN COMPLEX ENVIRONMENTS: ATTENTIVE TRACKING AND VIEW RECOGNITION

by

Jacob Isbell

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Timothy K Horiuchi, Chair
Professor Pamela Abshire
Professor Behtash Babadi
Professor Cornelia Fermüller
Professor Daniel Butts

Table of Contents

List of Figures

Chapter 1: Introduction

Bat echolocation is the unusual ability by bats to emit an ultrasonic sound pulse and measure the time until echoes begin to arrive (for estimating range) combined with the more general ability of mammals to determine the direction of sound. While senses like vision and hearing are important for our understanding of the human experience, it is not easy for us to understand what kind of world is experienced by the echolocation user. This work is motivated by a desire to better understand how echolocation creates a perceptual experience of the environment for the bat. It is focused on recreating possible biological processes used for echolocation, using robotics and computer processing to tackle the problems and scenarios echolocating bats would encounter on a daily basis. This ranges from tracking and capturing food (a lower level motor loop that takes precise timing and a quickly updating world model) to recognizing place and navigation, a higher level abstraction of environmental information gathered over a longer period of time. Sonar transducers can be used to mimic the echolocation ability of bats. A robotics platform in a laboratory provides the environment and object interactions needed to explore this phenomenon.

In addition to our motivations to ultimately model and understand the biological implementation of sonar-guided behavior, this work has applications for mobile, autonomous robotics. There are many circumstances when a drone may need to navigate in a dark building for stealth, through a building filled with smoke, through a forest with dense fog, or through tunnels filled with dust. Bats are well

adapted to these locations, often living in caves or dark, abandoned buildings [1]. Since standard cameras and LIDAR do not work well in these environments, sonar is a reasonable alternative or complementary sensor. Current laser and radar systems consume much more energy than a sonar system; this would reduce the robot's field time and potential range. The weight and cost of LIDAR systems can also reduce their feasibility of use [2]. Sonar has long been used for obstacle avoidance [3]–[5] Currently, the most familiar use of sonar systems is underwater. Since the speed of sound is much faster underwater, the effective range of sonar is greatly increased underwater. In the air, lightweight sonar systems are a good match for UAVs. One can imagine a lightweight, flying drone that can quickly maneuver through a dark house and provide a map based more on sensory features and not metrical details, closer to the way humans communicate with each other.

The ultrasonic frequencies used by bats (20kHz to 100kHz) are difficult to detect by most animals and have short wavelengths (~ 3mm to 17mm) that produce detectable echoes from small insects [6]. To localize the direction of echoes, bats (e.g., the big brown bat) have been shown to rely primarily on the use of interaural level differences produced by the head and pinnae, a common strategy for small mammals. The use of ultrasonic frequencies and a small head size strongly limit the use of phase-locking and interaural-timing cues for localization. To estimate range, the bat measures the time-of-flight of the echo from an emitted sound. From an auditory processing point of view, echolocation is unique in that the sound being analyzed is generated by the bat and is therefore both known and under the control of the bat. It is well known that bats change both the properties of the echolocation

pulse and the timing of pulses in response to their environment [7]–[9], but seldom has this dynamic behavior been adopted in artificial sonar systems.

1.1 Clutter and Aliasing

A typical operational assumption in echolocation is that all of the sounds following an emitted pulse are echoes from the most recent outgoing pulse. The duration of perceptible echoes resulting from a given pulse depends on the properties of the outgoing pulse (such as the amplitude, spectrum, and duration) as well as the properties of the environment (such as the distance, size, shape, orientation, and overall configuration of objects). A common-sense rule is that the next pulse should not be emitted until all perceptible echoes from the previous pulse have died out. In the majority of situations, bats appear to avoid this pulse-echo ambiguity, or "aliasing". Studies of big brown bats navigating in extremely cluttered environments, however, show cases where bats appear to tolerate such aliasing to sample the environment at a high-rate [7].

In close-quarters maneuvering, a high sampling rate is desirable when the angle to nearby objects is changing rapidly. Little is known about what bats do when a high pulse rate is needed to maneuver near objects in an environment that produces long-delay echoes, a situation that produces echo aliasing. Big brown bats have been shown to alternate between pulsing rapidly and pulsing slowly. Pulsing rapidly gives a clearer picture for close ranges while pulsing slowly gives a clearer picture for long ranges [7]. Dolphins also have been observed to increase their echolocation rate during prey capture [10], suggesting that this behavior is a result of the mechanics of

echolocation, not necessarily the specific environment or animals involved. Bats and Dolphins are believed to have evolved the skill of echolocation completely separately, a phenomenon known as convergent evolution [11]. A possible strategy for dealing with these long-delay echoes might be to reduce the intensity of the call or reduce the low-frequency components of the chirp to reduce the distance over which the perceptible acoustic pulse travels. Bats have also been observed to change the spectral content of consecutive pulses, largely by shifting the entire pulse up or down in frequency. The spectral signature of the returned echoes can then be used to assign them to a specific pulse [8], [12]. This technique can also help bats when hunting in groups, where distinguishing between the multiple calls and echoes occurring from different bats can be difficult [13]. One strategy bats have for dealing with cluttered environments is to use spectral changes in the reflected echoes to distinguish objects [14]. Object qualities such as hardness, texture, and motion will affect the spectral change of an echo from its incoming sound. *Carollia perspicillata* has a unique method for shaping the beam of their echolocation. They emit sound from two nostril holes, which creates a phasing pattern the changes depending on the frequency used, allowing the bat some control over the shape of the echolocation beam [15].

Radar has many parallels to sonar and bat echolocation; both are active sensing and use reflections to gain information about objects. Some techniques used in radar are very similar to those used by bats. For example, changing the emitted frequency allows radar systems to increase the effective sampling rate [16], [17]. Another technique utilized by radar systems is to transmit multiple pulses in a short temporal pattern (or “code”). Different codes can then be used to identify different

pulses [17], [18]. When the task is to track a specific target object (e.g., an obstacle the bat is maneuvering around), an attentional mechanism can be used to ignore the background and any aliasing that may be occurring. This approach works well until an 'aliased' echo arrives at or near the time of the tracked echo. Parallel to the *Carollia perspicillata*, phased arrays have been used in radar to control beam shape and directivity.

1.2 Hippocampal Place Cells

The hippocampal formation in the mammalian brain is well known for its population of 'place cells', a type of neuron that responds when an animal is in a particular place in its environment. Studies in the rat suggest that these cells use internal odometry signals (allowing the system to operate in darkness) as well as external sensory cues (allowing the system to recognize places and correct the odometry system) [19]. These cells have been found in rats, bats, monkeys and humans [20]–[22]. In the flying, echolocating bat, neurons with very similar properties have been found [23]–[25]. Unlike rats, bats have the uncommon ability to perceive the three-dimensional locations of objects by actively emitting sounds and localizing the reflections [26], allowing the bat to navigate where other sensory systems, such as vision, are ineffective. Although the signal processing and neural mechanisms with which bats recognize places is still largely unknown, modeling this capability with biologically-plausible sensors and robotics can give us insights into problems that bats encounter and motivate future behavioral and neurophysiological experiments with bats.

In robotics, it is common to represent location with a metrically precise coordinate. Systems like GPS provide an estimate of a point in space. When maps are created, they are often in precise geometrical coordinates, using precise distances and measurements. In contrast, place cells are active over a range of space. Many questions come up when place cells are modeled that do not come up when conventional coordinate systems are used. How are the centers of activation chosen? How broad of an area should a place cell be active over? How densely should place cells populate a given area? If we think in terms of how an animal functions, these questions can be posed in relation to completing tasks. Centers of activation may be chosen that correlate with different tasks, the area of a place cell could relate to the area over which a task can be completed, the complexity of the environment could determine how densely place cells populate a given area. Navigational maps created using place cells may have a more functional representation than a precise one; sequences of place cell activation may represent paths from one place to another [27], [28]. Intuitively, we have an idea of places at very different sizes and scales. Our desk is a place, our house is a place, our city is a place. In different contexts (depending on the task), it is useful to have different scales of place. Animals have place cells that activate over different scales that relate to the kinds of tasks and environment they are in [25]. There may be smaller, more densely populated place cells in an area with more local activity, like a nest or foraging ground, while areas like long paths may have larger place cells.

Groups of cells with different characteristics have been found in the hippocampus. In the entorhinal cortex there are cells that are active along boundaries

and walls called border cells [29], [30]. There are also head direction cells that fire in relation to the direction of an animal's head relative to its environment [31]. These cells fire irrespective of what place the animal is in and complement the functionality of place cells. Their behavior is similar to a compass in that they both respond to a global direction. Both head direction cells and place cells may contribute to an animal's navigational ability. The entorhinal cortex is also home to grid cells. These cells fire repeatedly at regular spatial intervals, creating a triangular grid with activation at each of the vertices [32]. While the role of grid cells is not completely understood, they may contribute to an animal's sense of odometry and path integration [20]. Others have posited that they may be a part of an error correction system that adds robustness to noise [33]. There are even cells that have characteristics of both head direction cells and grid cells [34].

Another type of cell (of particular interest to our work) found in the primate hippocampus is the spatial-view cell [22], [35]. This type of cell responds when an animal is looking at a certain view, regardless of place, head direction, or eye position. This cell will still be active if the view becomes obscured or lights removed, as long as the eyes look in the direction that the view was previously seen. More than recognizing place, these cells will recognize objects or landmarks in the environment and their allocentric position. This may help animals determine what tasks may be relevant within eyesight rather than at the specific location of the animal. While recognizing views and recognizing place are not exactly the same, they are strongly related. These view cells may contribute to the activation of place cells.

For the flying bat, a two dimensional map is not sufficient to represent the animal's movement. Place cells have been observed in the bat that correspond with three dimensional volumes in space [24]. There are also head direction cells that are tuned for positioning in three dimensions as well [36], giving the bat the ability to navigate through complex environments with three dimensional features, such as cave systems and cliff sides. Goal-Vector cells have been found in the bat that relate the distance and direction to certain goal points [37].

1.3 Similar Robotics Work

Although most robotic explorations into mapping and navigation have focused on variants of the SLAM (simultaneous localization and mapping) algorithm using light-based sensors (e.g., computer vision or LIDAR) [38], [39] for metrically-accurate maps, little work has been done exploring how a bat might use sonar to accomplish the same task. One good example is that was BatSLAM [40], a biomimetic sonar system that used odometry and sonar to map an area of their laboratory. Because odometry is quite inaccurate due to wheel slippage and other errors, such as compounding inaccuracies in estimating direction and position, sonar was used to provide error correction. Their system first drew paths of motion based solely on odometry. When the sonar-based recognition system recognized the current location from a prior visit, it updated the odometry system to match its memory and propagated the correction to earlier time steps for consistency. This was sufficient to correctly create a map of the area with little error. While this approach showed that sonar was able to aid place recognition, it did not do so in a biologically-plausible manner. Over the robot's path, 6000 sonar measurements were taken, and 3300

different places were established. While this system provides a method to maintain an estimate of the robot's position, it does not seem to reflect what little is known about how biological memories of the environment. Memorizing 3300 different places all within one environment is computationally and memory-intensive; it is not a biologically-plausible algorithm. While our study attempted to show that odometry is not needed for view recognition, incorporating odometric information can provide a strong framework for unsupervised mapping. For example, a new 'place' can be created when a system, using odometry, estimates it is a certain distance from any other 'place'. The work presented here addresses the question whether echo view cells can be recognized over an extended region using only a narrow-band (~40kHz) sonar in a laboratory environment. Unlike the place cells that signal when the animal is in a particular area (i.e., the "place field") based on a combination of odometry and sensory inputs, we are constructing "echo view cells" that recognize previously encountered views (i.e., an 'echo fingerprint') based solely on sonar. Phenomenologically similar to primate "spatial view cells" that are active when the animal is gazing at a particular set of objects (over a limited field-of-view) [35], these echo view cells recognize previously memorized echo patterns. Unlike primate spatial view cells, however, object range is included in the pattern and thus the echo view cells fire over a small region of the environment.

Another recent paper explored the idea of recognizing place with sonar in three different locations [41]. Using a very precise sonar sensor they measured the echo response at positions over a wide range of angles and along a linear, 10m long path. They collected an enormous amount of data (over 20,000 echo traces) and

evaluated whether the echoes varied smoothly over angle and distance as well as whether unique locations could be classified. Most of the data came from angular variation; large translational steps contrast the high angular resolution. They also found places that were difficult to distinguish between, mainly in open areas with few objects to sense, but concluded that sonar is enough to recognize most locations. To compare sonar traces they used the Euclidean distance between two sonar traces. To estimate a range over which a place can be recognized, the catchment distance was used, a tool borrowed from visual place recognition. When they were comparing different positions along a linear path, they compared the same precise angle (0.1 degree error) from the different positions. This is much more precise than an animal can hope to achieve, in reality both angle and position will be changing at the same time. We have shown in this study how sensitive an echo signature can be to changes in angle; we expect place recognition to be tolerant to moderate changes in the sensing direction. Our study can complement this one by providing a wider, two-dimensional range of positions for comparison as well as removing the need for very precise angular measurements.

Another example of a sonar system used for map creation is the ‘Robat’ [3]. Here, the precise position of echoes ensonified by sonar were recorded and added to a metrically precise map. A biologically plausible sonar using one emitter to mimic a bats mouth and two receivers to mimic bats’ ears was used. The Robat autonomously navigated a greenhouse environment on its own. The map they created consisted of boundaries encountered rather than specific objects themselves; this approach makes it useful for obstacle avoidance. Every echo was considered to be an obstacle. In a

more complex environment, many echoes can be created from things as simple as a branch on the ground. In reality, not every echo will be an obstacle. A very interesting part of this study that works towards addressing this issue was the attempt to classify echoes as plants or not plants using a neural network. They achieved a 68% accuracy (where chance would have given them %50). A major difference between this work and ours is that they were classifying individual echoes based on a temporal-multiband signature, where we classify a view based on the many echoes it is composed of with each echo represented simply by its magnitude. Being able to classify objects using sonar would be a large step in determining appropriate actions to take given a sonar image.

Chapter 2: Managing Clutter in a High Pulse Rate Echolocation System

2.1 Materials and Methods

2.1.1 Hardware

The sonar system used in the work presented here consists of three custom-modified MaxBotix® sonar transducers, similar to the MaxBotix XL-MaxSonar®-EZ™ commercial series of sonar range finders, a custom PIC® 18F2620 (Microchip Technologies Inc.) microcontroller-based sonar controller board, a Futaba S148 hobby servo, and a computer interface to both record and display echo signals and control the servo to orient the sonar (shown in Figure 2.1). The transducers act as both a speaker and a microphone. They resonate at 40 kHz and will only detect signals near this frequency. The custom sonar boards report a logarithmically-compressed envelope signal as an analog voltage. This compression allows the output to report the very wide dynamic range of amplitudes that occurs with sonar without saturating. The maximum working range of this sonar is 7.65 meters. These transducers were custom modified to provide more control over the timing and duration of the outgoing pulse, a louder outgoing pulse, and access to a log-compressed envelope of the transducer response. All these functionalities are now commercially available through MaxBotix. The transducers are placed in a 3-D printed mount on the servo motor. In this demonstration system, the transducers

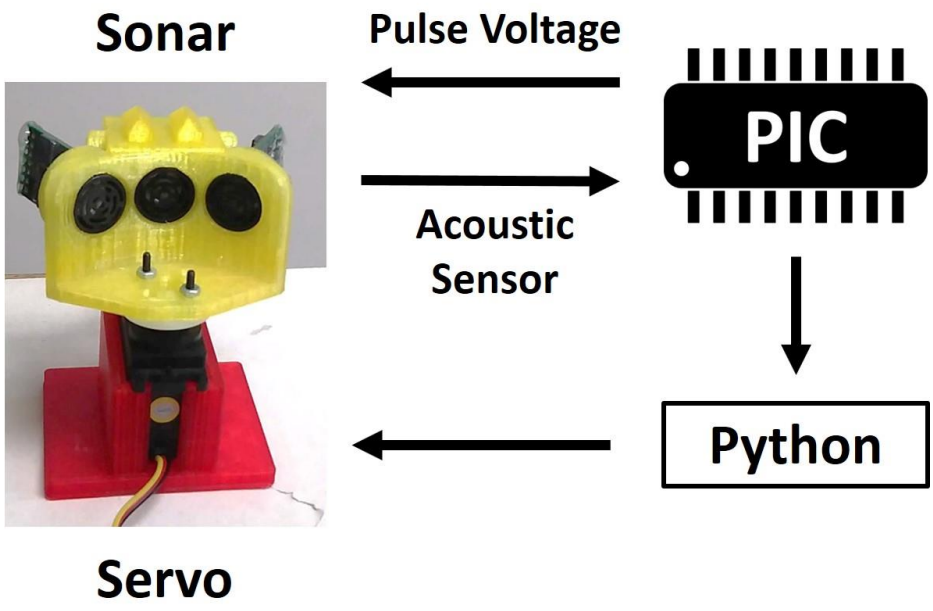

Figure 2.1: The flow of signals through the hardware. The microcontroller sends pulse voltages to the transducers and reads the acoustic voltage off the transducers. This data is sent to Python on a PC, which also controls the servo motor.

transmit and receive over a cone of about +/- 30 degrees (-6dB beam width), so the transducers are held facing thirty degrees apart to ensure sufficient overlap and coverage of the area in front of the transducers for binaural localization based on interaural level differences. The ultrasonic pulse trigger-timing and analog-to-digital (A/D) conversion is done by the microcontroller. The majority of the data processing is performed on the microcontroller to ensure a quick response.

The sonar system executes four steps: pulsing, sampling, processing, and communicating. A short duration pulse voltage (~0.25 ms) is supplied to the transducer, however, due to the resonant quality of the transducer, the emitted sound has a ring-up and ring-down period, resulting in an extended pulse duration of about 1ms. Following the pulse, the transmitting transducer continues to ring for several

milliseconds. Echoes can be detected during this ringing period once the amplitude has diminished sufficiently, so a short two millisecond delay is incorporated before sampling begins. The log-compressed envelope voltage is sampled every eighth of a millisecond, a sampling frequency of 8kHz. An object is detected when the temporal derivative of the envelope switches from positive to negative, denoting a peak. The range is determined by finding the time when the envelope reaches its peak value. Envelope voltages on all transducers are recorded at the time of the peak. Our sampling time of an eighth of a millisecond gives us a range resolution of 2.14cm or .84in. We sample for 255 time bins, giving us a range of 5.5m or 18ft. Following the sampling period, echo data is transferred via serial interface to a PC and all further processing on the information is performed on the PC. An example of this data is shown in Figure 2.2.

2.1.2 The Tracking Cycle

The sonar system executes four repeated steps: pulsing, sampling, processing, and communicating. As part of the cycle, there is an added delay interval that is used to reject aliased echoes (discussed in section 2.1.5). A few of these steps are shown in Figure 2.2 for two cycles. In these examples, a short duration ultrasonic command pulse (~0.25 ms) is used, however, due to the resonant quality of the transducer, the duration of the acoustic pulse is extended. Following the pulse, the transducer continues to ring for several milliseconds. Although echoes can be detected during this ringing period, their amplitudes are difficult to estimate, so a short two millisecond delay (i.e., dead-zone) is incorporated before sampling begins. The log-

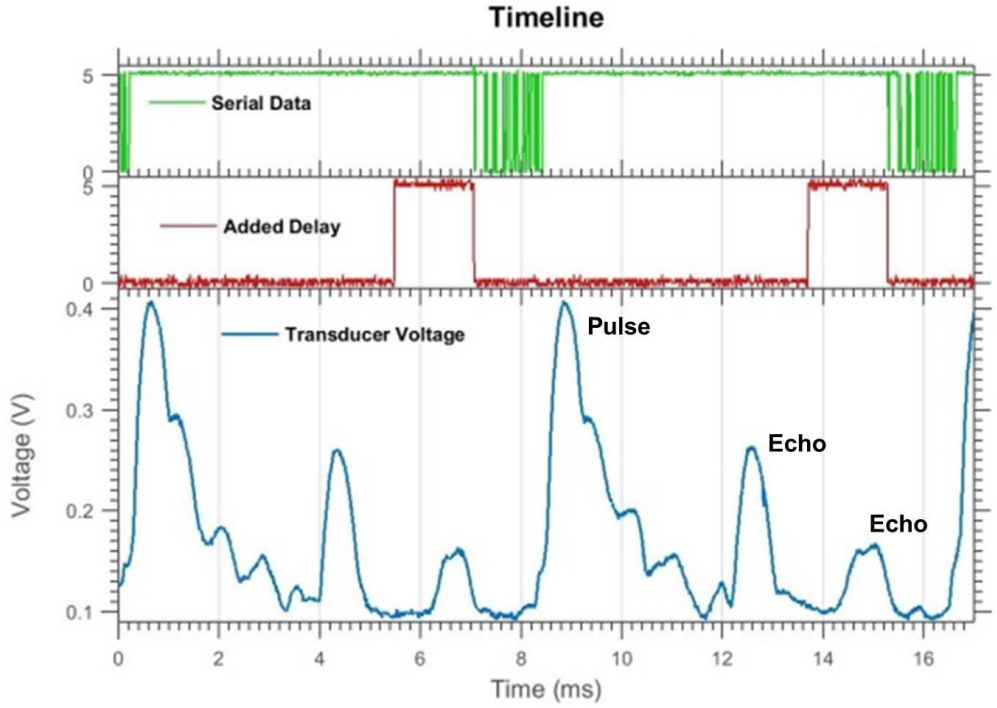

Figure 2.2: An oscilloscope readout of two pulse-echo cycles (without aliasing) showing the transducer envelope voltage (bottom), serial data transfer (top), and added delay (middle). The added delay flag is set high when the delay is occurring. Objects can be seen as distinct peaks in the transducer voltage trace. Pulsing and sampling the transducer takes 5ms, then there is a 1.5ms delay and 1.5ms of serial data transfer. The whole cycle takes about 9ms.

compressed envelope voltage is sampled every eighth of a millisecond. Object detection begins when the temporal derivative of the envelope exceeds a threshold of approximately 3.4dB over an eighth of a millisecond. Once the peak of the envelope has been reached, the object range is determined by the time since emission and the direction is estimated using the amplitudes on the two transducers.

At low pulse rates, the echoes are monitored for a period of time associated with the maximum range of the sonar and an extra delay would be added after

transmitting the recorded data. In the case of fast pulsing where a target is being tracked, once the target echo is received, a short data burst is transmitted and the next cycle is initiated. After detecting the target echo, the tracking window (in time) is updated and the intensities of both transducers are compared to rotate the servo motor to center the target echo. At this time, temporal windows before and after the target echo are monitored to detect if other echoes are about to overlap with the target echo. This information is used to initiate the various reactive strategies to avoid interference with target tracking (described in section 2.1.5-2.1.7).

2.1.3 Target Tracking

There are occasions when the echo from the target disappears completely due to interference or occlusion by an object in the foreground. The tracker continues to search for the target at the same range for up to three cycles after the object disappears. If the object does not reappear, it will begin looking for a new target at a pre-specified acquisition range.

For the purposes of this study, the tracker is programmed to initially find the target at a single, pre-specified range (about 33 cm) and then follow it in range and in the horizontal plane by turning the sensor head to center the object. Centering is accomplished by rotating the sensor head until the detected amplitudes of the target in the two transducers are approximately equal. Only horizontal angles are considered. Since the echo amplitude is logarithmically compressed, the difference between left and right outputs corresponds to a ratio of the two received amplitudes. This ratio (invariant to echo amplitude) can be mapped to a specific angle. This mapping is

defined by the spatial sensitivity and placement angles of the receivers and is found empirically. The ratio is monotonic and allows for reasonable angle measurements over a range of +/- 30 degrees. Outside of this region, only one transducer will produce a significant response, allowing only a coarse approximation of direction. The response of our system at various angles is shown in the top graph of Figure 2.7. The range of objects is determined by the time when the echo is received (i.e., time-of-flight). In practice, this is a very stable measurement that is minimally affected by noise. The echo amplitude, however, is very sensitive to factors such as the shape and orientation of the target, interfering reflections and echoes, and positioning of the transducers. At high repetition rates, a reverberant room can become filled with sound, introducing significant background interference. To avoid wild oscillations in the servo motor pointing, the system is restricted to moving a maximum step of 5 degrees between echoes.

Once an object is found at the pre-specified acquisition range, it is labeled as the target and tracked. In the next pulse cycle, the sonar will expect to receive an echo within 6.3 cm of the previous target range. By restricting the temporal size of the tracking box, all echoes other than the target are ignored allowing the system to track a single object in the midst of other objects. Analog-to-digital sampling is performed with a period of an eighth of a millisecond and thus the range resolution is 2.1 cm/sample.

2.1.4 Aliasing and Clutter

In the rapid pulse mode, the maximum detection range for the sonar system is limited by the interpulse interval. If an object has an echo time that is greater than one

pulse period, it is detected by the system in the next pulse cycle. It is then perceived as having an echo time that is one pulse period less than it actually is. Since this distortion is caused by sampling related to each pulse, we call it aliasing. This is demonstrated in Figures 2.3 and 2.4. While the perceived direction of this "phantom" object is unchanged, the range is wildly incorrect and may even overlap the echo from the tracked target. The techniques presented in this paper aim to keep the range and angle measurements of the target clean. This can be done by keeping other echoes far enough away (in time) to not overlap the target echo (~.5 ms). If that is not possible, the goal is to reduce the amplitude of the obstructing echo as much as possible.

Two strategies specific to problem of aliased echoes overlapping the target echo are presented: First, by using an adaptive delay, the interpulse interval can be manipulated to change the relative time of the aliased echo. This changes the perceived range of the alias to prevent it from overlapping with the target. Second, the sonar system can use movement to prevent objects in the background from falling in the main path of the sonar beam. This reduces the magnitude of clutter echoes. These strategies may not always work, particularly if the aliased object is close in range to the target and the sonar beam is too wide for the movement strategy to avoid illuminating the aliasing object. In this case, beam forming of the transmitted pulse by firing both transducers in a phased manner can be used to increase the amplitude of the target echo and decrease the amplitude of the aliased object echo. This can

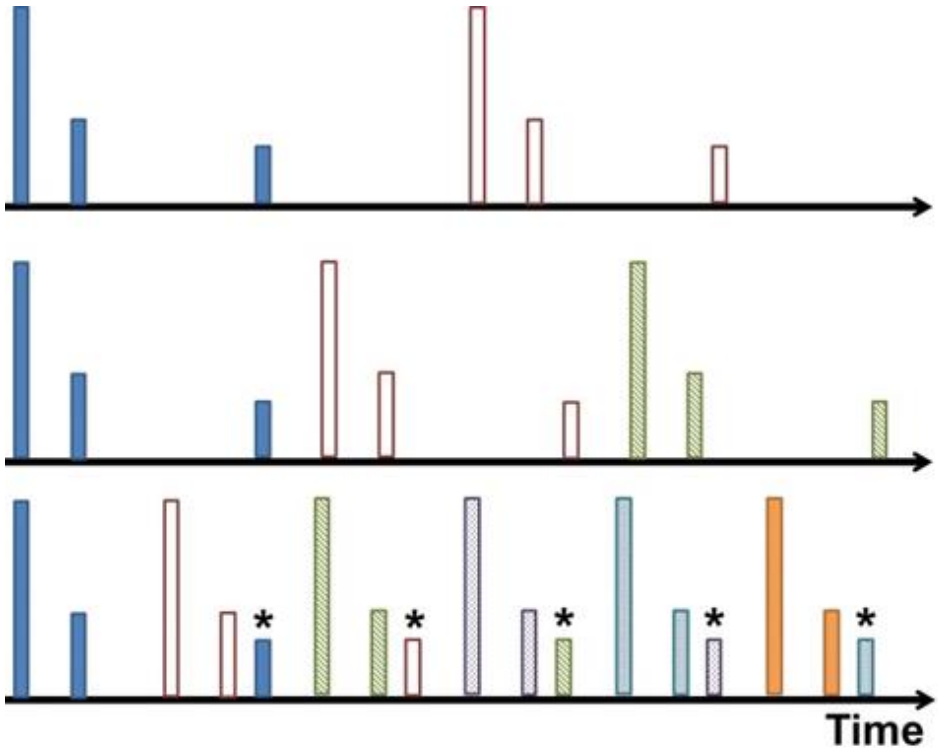

Figure 2.3: Aliasing visualized. In this cartoon example, each timeline has pulses (represented by tall lines) and received echoes (represented by shorter lines). Each pulse and its echoes are given a unique color. From top to bottom, the interpulse interval decreases until a new pulse occurs before all echoes from the previous pulse are received, shown in the bottom timeline. The echo is misinterpreted as a closer object associated with the latest pulse. This is the aliased echo and is labelled with an asterisk.

also be effective in non-aliasing situations where a distractor object at the same range (but different angle) is causing interference.

2.1.5 Method: Adaptive Delay

The range at which the aliased echo appears is dependent on the time between sending pulses. To control this, a variable delay period is inserted before sending the next pulse. Increasing this delay shortens the aliased echo time, making it appear to

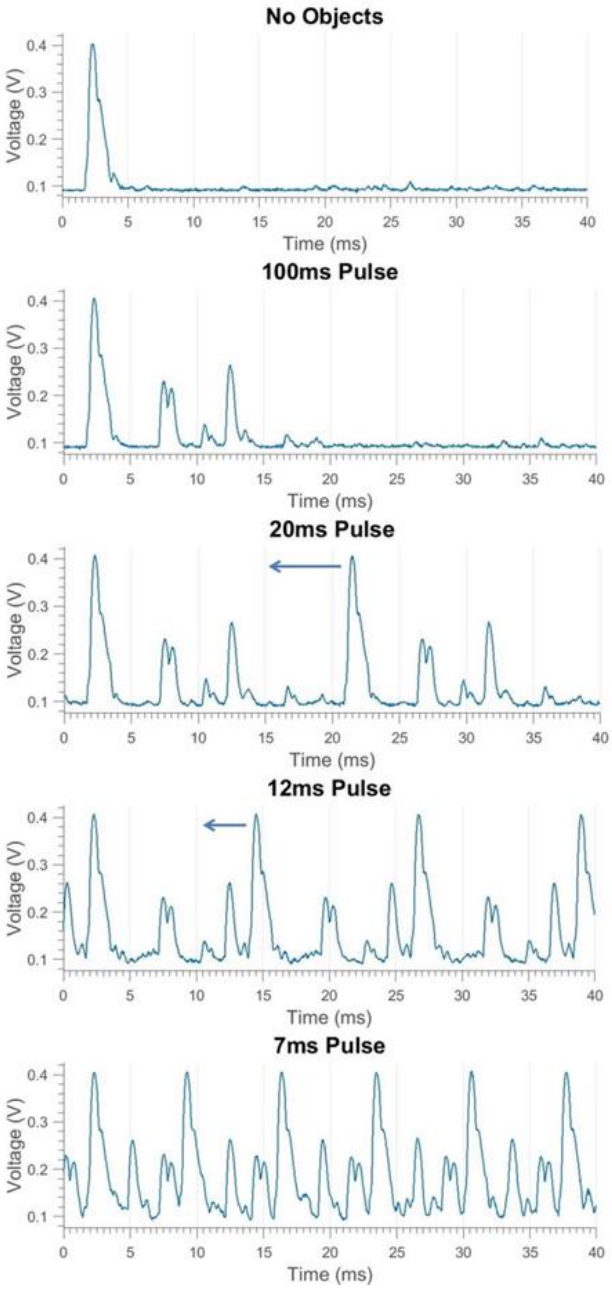

Figure 2.4: Transducer envelope of pulses and echoes at different repetition rates demonstrating aliasing in the bottom graph. The outgoing pulse peaks at 0.4V, overlapping echoes from two closely-spaced PVC pipes are seen peaking at 0.2V, and a single loud echo made by a square poster board is seen peaking at 0.25V. The interpulse interval is decreased in each graph until a new pulse occurs before all echoes from the previous pulse are received, causing an aliasing condition where the poster board incorrectly appears at short range.

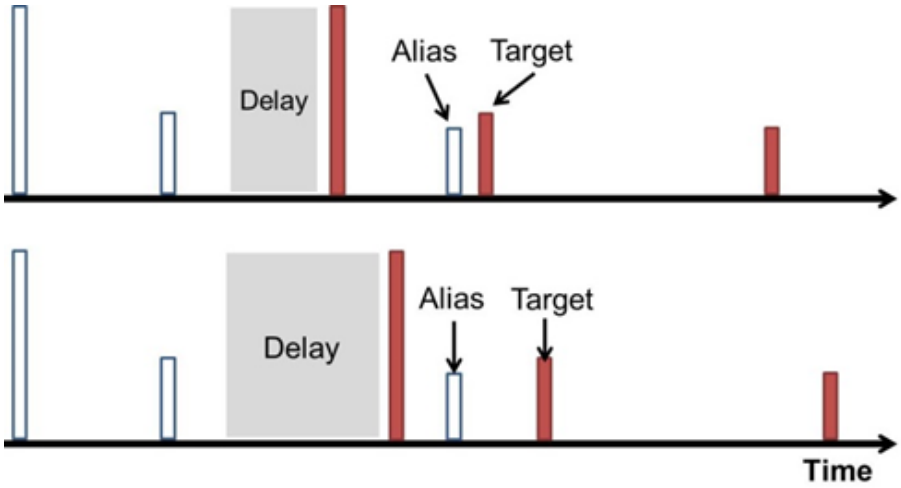

Figure 2.5: Manipulating the received time of an aliased echo. The tall line represents the pulse and the short lines represent the echoes. The echoes associated with a given pulse are the same color. The top timeline shows an alias (white) that is close to interfering with the first dark echo, the target. The introduced delay is increased (in the bottom timeline) to shift this aliased echo away from the target in time. Similarly, an alias on the other side of the target can also be shifted away by decreasing the delay (not shown).

move closer to the sonar. Decreasing the delay increases the aliased echo time, making it appear to move away from the sonar (an example is shown in Figure 2.5). The alias rejection system introduces a delay interval with a maximum of three milliseconds into the timeline. The interval length is changed in eighth millisecond increments based on where the aliased echo appears relative to the tracked target. If an aliased echo is within 5 range samples, or 10.7cm, of the target echo, the delay interval will be changed to repel the aliased echo. For an aliased echo that appears closer than the target echo (i.e., in between the target and the sonar system), we increase the delay to move the aliased echo away from the target echo; an aliased echo further away than the target echo decreases the delay. If the delay reaches its maximum amount or if it is decreased to zero, the delay value is reset to 1.5

milliseconds (half of its maximum value). This will cause an aliased echo to jump to the other side of the target echo, being shifted by 12 range samples. If there is an aliased echo detected on both sides of the target, the delay is shifted by a large amount, equivalent to 11 range samples, in an attempt to clear both aliased echoes away from the target echo. This process is summarized below.

If alias in front

    Increase delay

If alias in back

    Decrease delay

If alias in front **and** alias in back

    Large delay **shift**

If delay is minimum **or** delay is maximum

    Reset delay

While we have assumed a relatively isolated target object to track, a real second object in close proximity to the target cannot be "rejected". In this case, the alias rejection system would continuously shift the delay, resulting in oscillations of the delay shifting and resetting when the delay interval reaches its limits. To prevent these oscillations, additional code is used to recognize authentic (i.e., non-aliased) echoes.

The most notable difference between an authentic echo and an aliased echo is their reaction to a large shift in the interpulse interval, a delay jump. An alias will be moved a significant amount, while an authentic echo will not be moved at all.

Although a real object can still move noticeably, at low speeds (< 3m/s) it will not jump more the one range sample at a time.

The alias rejection system makes large delay shifts in three different scenarios: when the delay interval reaches its maximum, its minimum, and when two aliases sandwich the target (one on either side). The system uses these events as triggers to look for an authentic echo that remains in the same location. This is especially appropriate since an authentic echo triggers an oscillation that causes the delay to jump when the interval reaches a maximum or minimum. If an object does not move after a delay jump, it is recognized as an authentic echo and will not activate the alias rejection system. This is similar to a technique used in radar where a map of stationary clutter is memorized and removed [17].

2.1.6 Method: Movement

An alternative method to avoid sonar aliasing is to reposition the sonar beam such that objects in the background do not generate echoes. The effectiveness of this technique will depend on using a relatively narrow transmission beam. Depending on the species of bat, transmission beam widths can range from 22 to 90 degrees [10], [18], [42]. The sonar beam width used in this study is approximately 30 degrees. When the sonar moves around, different sides of objects are exposed to the sonar. In general, this will complicate a decision to change the sensing angle, since the acoustic properties of an object can change greatly from different perspectives. To demonstrate this, two different objects were used as the aliasing object in two different trials: a large 46cm (1.5ft) diameter cardboard tube and a 30cm wide, open cardboard box. The sonar was moved around a target object to continue aiming the

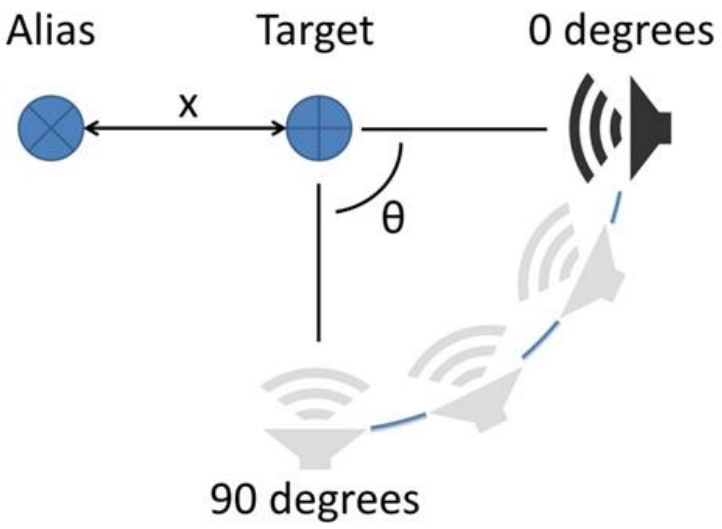

Figure 2.6: Alias rejection via movement. The sonar system (speaker) is kept a constant distance from the target. The alias is located at a distance x from the target. The sonar is rotated around the target by angle ɵ to shift the view of the system.

beam at the target at the same range but resulting in different backgrounds (Figure 2.6). As the sonar moves, the transmission beam is moved away from the aliasing object and the magnitude of its echo decreases. Theta is the angle of rotation the sonar system has made around the target relative to its starting location. For this study, only one transducer was used.

2.1.7 Method: Beam Shaping

A third strategy for reducing the effect of aliasing and clutter objects is to shape the acoustic beam so that only the target object is ensonified. With the two-transducer system used in the study, this is performed by transmitting with both transducers to create an interference pattern that has peaks and nulls that can be used to reduce interference. Plots of the beam shape are shown for a single transducer, the two transducers firing synchronously, and the two transducers firing out-of-phase

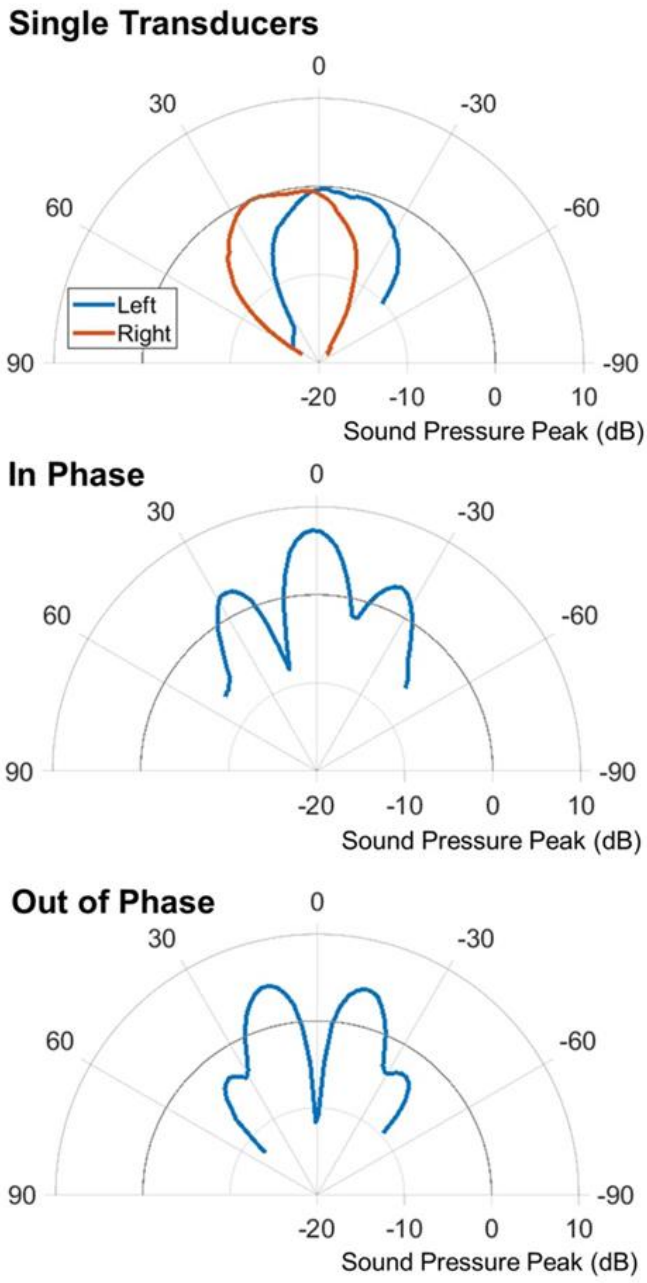

Figure 2.7: Polar plots of the different firing patterns. Top shows single transducer pulses from the left and right transducers. Middle shows the synchronous in-phase firing pattern. Bottom shows the synchronous out-of-phase firing pattern.

(Figure 2.7). The synchronous in-phase firing pattern has a loud frontal lobe that is relatively narrow with weaker lobes on either side. The -6db width of the front lobe is 19 degrees (compared to 62 degrees of a single transducer alone). The stronger,

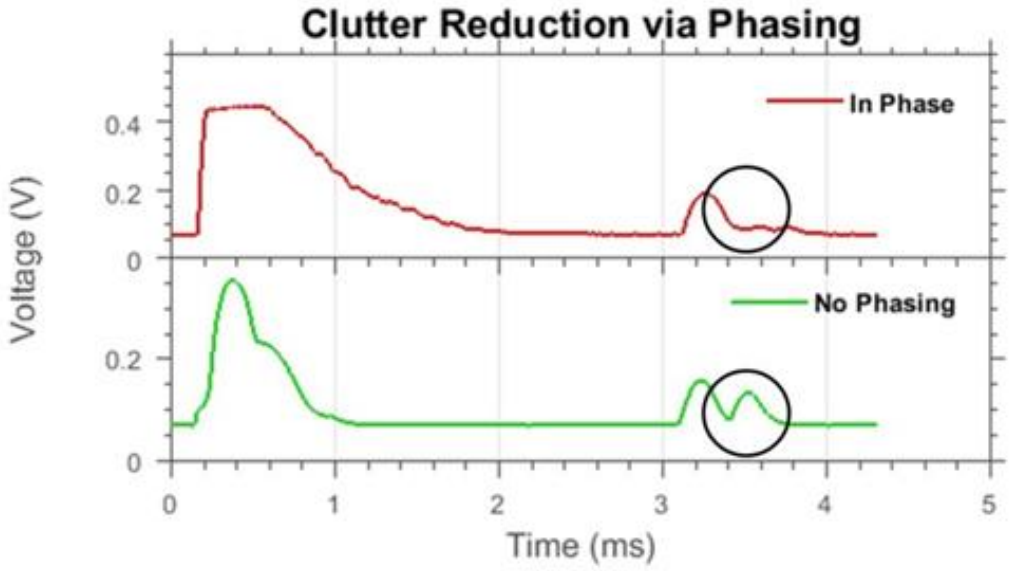

Figure 2.8: A best-case example of clutter reduction using beam shaping. Shown are two echo traces from the same scene with different beam shapes. Two objects are present, the first echo is the target object (~3.2 ms); the second echo is from the clutter object (~3.5 ms) which is circled. When in-phase firing is used, the clutter echo is greatly reduced in amplitude.

narrower central lobe would allow more precise ensonification of a target while reducing echoes from other directions. It is important to note that the patterns presented here represent the transmitted beam only. The sonar hardware presented here does not allow phased detection; although that is an additional capability in other systems that would further improve selectivity.

Figure 2.8 shows an example of how using this firing pattern can affect an echo trace. There are two objects in the field of view, both PVC pipes of equal

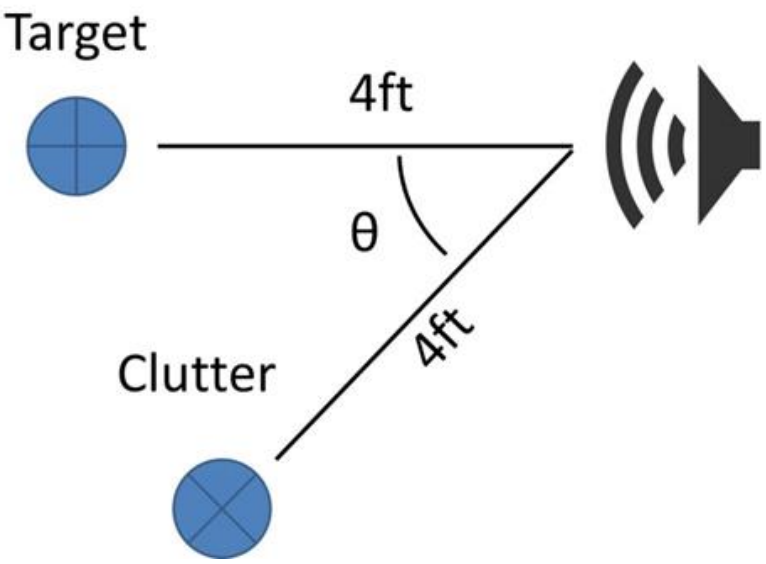

Figure 2.9: Clutter rejection using beam shaping. The sonar system faces a target that is 4ft away. The clutter object is also four feet from the sonar but is rotated around the sonar system, changing its angle in the view of the sonar.

diameter. Using only a single transducer without phasing, the clutter echo impinges on the target echo and disrupts the information conveyed. Using two transducers fired synchronously, the cluttered object has a significantly reduced magnitude and the target will not be as affected by the clutter. This example, however, represents a best-case scenario where the clutter object falls in the low trough of the firing pattern. Utilizing this system in arbitrary object configurations is not trivial. Since angle estimation is very noisy in this sonar system, predicting the effects of beam shaping can be error-prone and not guaranteed to be beneficial.

The experimental configuration used is shown in Figure 2.9. With the target centered in the sonar view, the clutter object is moved to different angles relative to the center. Both objects are 5cm diameter PVC pipes at a range of 122cm (4ft).

2.2 Results

2.2.1 Adaptive Delay Results

Figures 2.10 through 2.12 show the system in action, presenting consecutive graphs in time that demonstrate system functionality. The three figures represent the three different cases of aliases: aliases moving toward the target from the front, aliases moving toward the target from the back, and two aliases sandwiching the target. In each case it is assumed that the target starts clear and unobscured. The adaptive delay prevents the target from becoming obscured in all the cases.

2.2.2 Movement Results

Figure 2.13 shows the results of the movement study with the aliasing object at 4, 5 and 6 feet (labeled x) away from the target. At the same angle, larger x values push the aliasing object farther away from the center of the beam and cause a larger reduction in magnitude. For the column, at a 60 degree angle of rotation, the aliased echo amplitude had been reduced to below 19% of the target echo amplitude for all distances of x. For the box at the same angle, the amplitude was only reduced to 57% of the target echo amplitude in the worst case (x=4 ft). This highlights the role of the object geometry. It should be noted that the measured amplitudes are logarithmically-compressed acoustic amplitudes and the actual percentage change seen will vary with signal level.

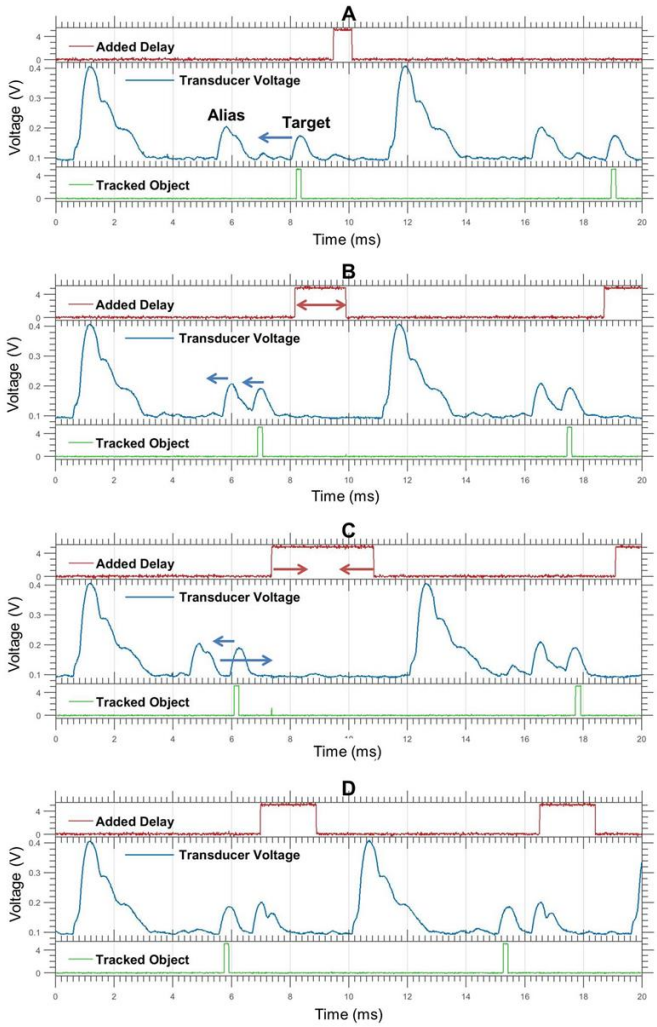

Figure 2.10: Oscilloscope showing transducer voltage, delay, and tracking for an approaching target. The added delay bit is high when the delay is occurring. The tracking bit is high when receiving the echo of the object being tracked. These graphs are a sequence of events in real time, from A to D. Only two significant objects are present, the target marked by the tracking bit, and the aliased echo. The only object moved was the target; the apparent movement of the alias is due to the delay change. The arrows show movement change for next frame. A to B-The target moves forward, toward the alias. B to C-The target continues forward, the alias is pushed forward by the increasing delay. The delay buffer becomes maximized. C to D-The delay buffer jumps down after reaching its maximum. This causes the alias to "jump" behind the target.

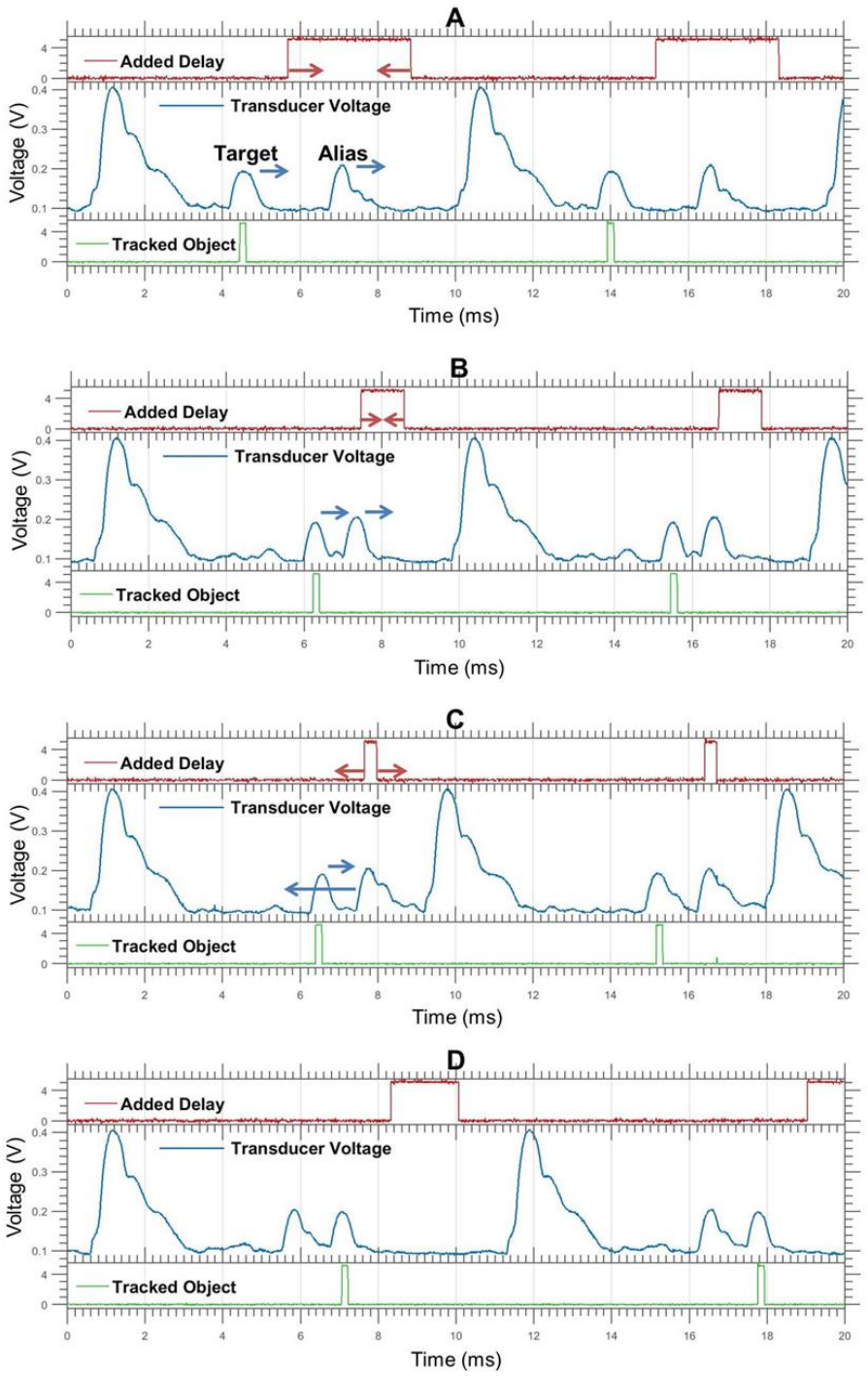

Figure 2.11: Continuation of figure 6 for a retreating target. A to B-The target moves back; the alias is pushed back by the decreasing delay. B to C-Both echoes continue backwards, the delay buffer reaches its minimum value. C to D-The delay buffer jumps upwards, causing the alias to jump forwards in front of the target.

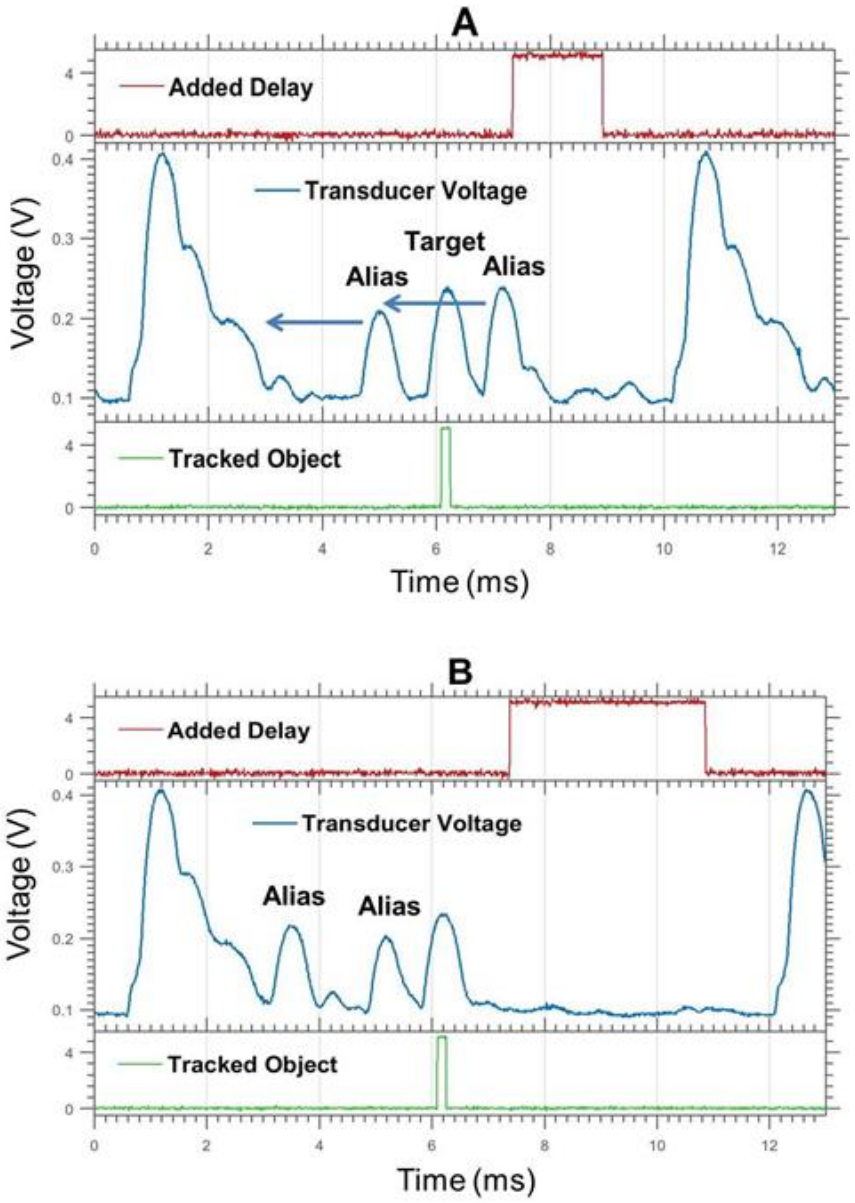

Figure 2.12: A target being sandwiched by aliased echoes. Here the alias in front of the stationary target approaches the target and triggers a delay jump. This clears both aliases away from the target, drawing them both forward of the target.

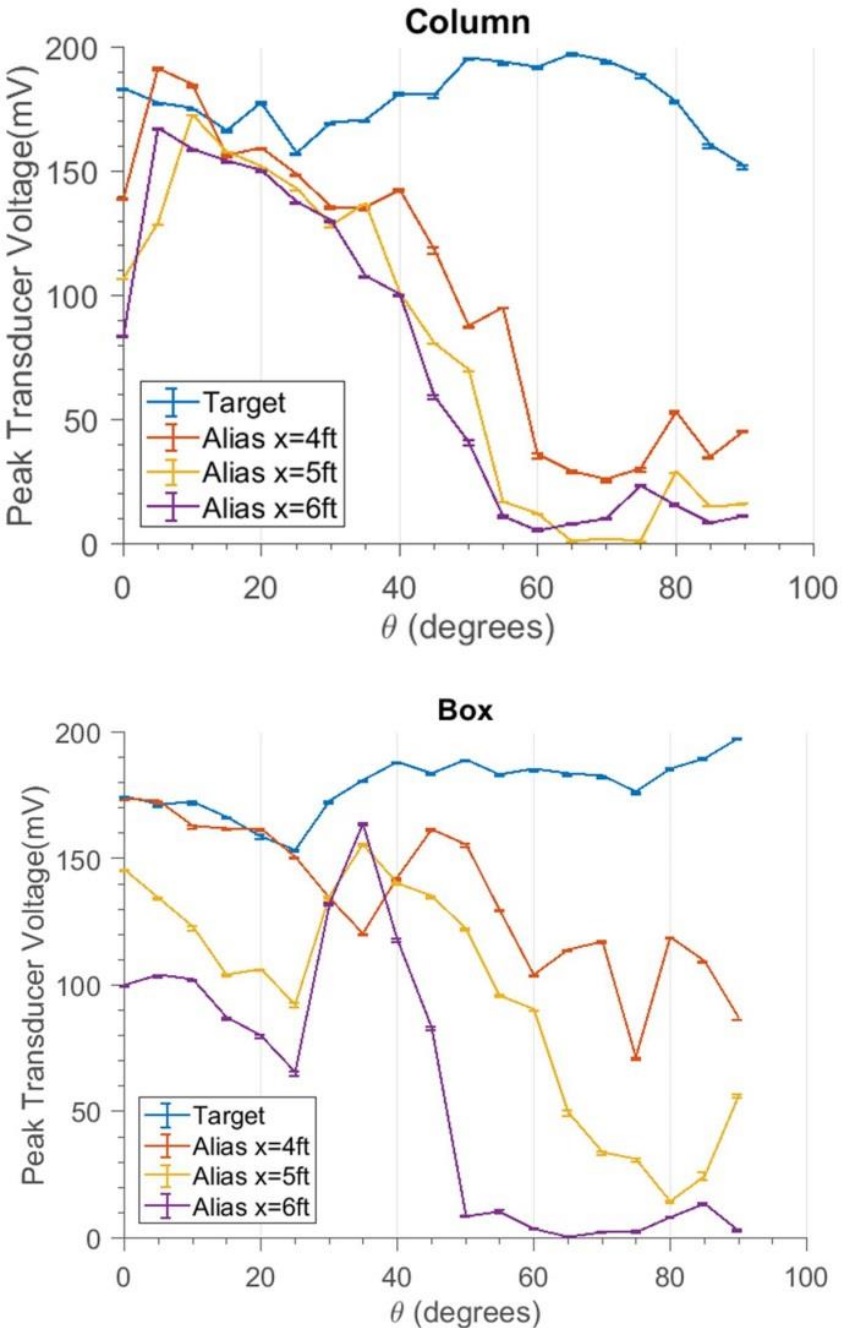

Figure 2.13: Traces showing the echo response of the target and the alias at different angles. The target trace (blue) gives a baseline for comparison. The 'Alias' traces reduce in amplitude as the angle increases. For larger distances x the amplitude decreases even more.

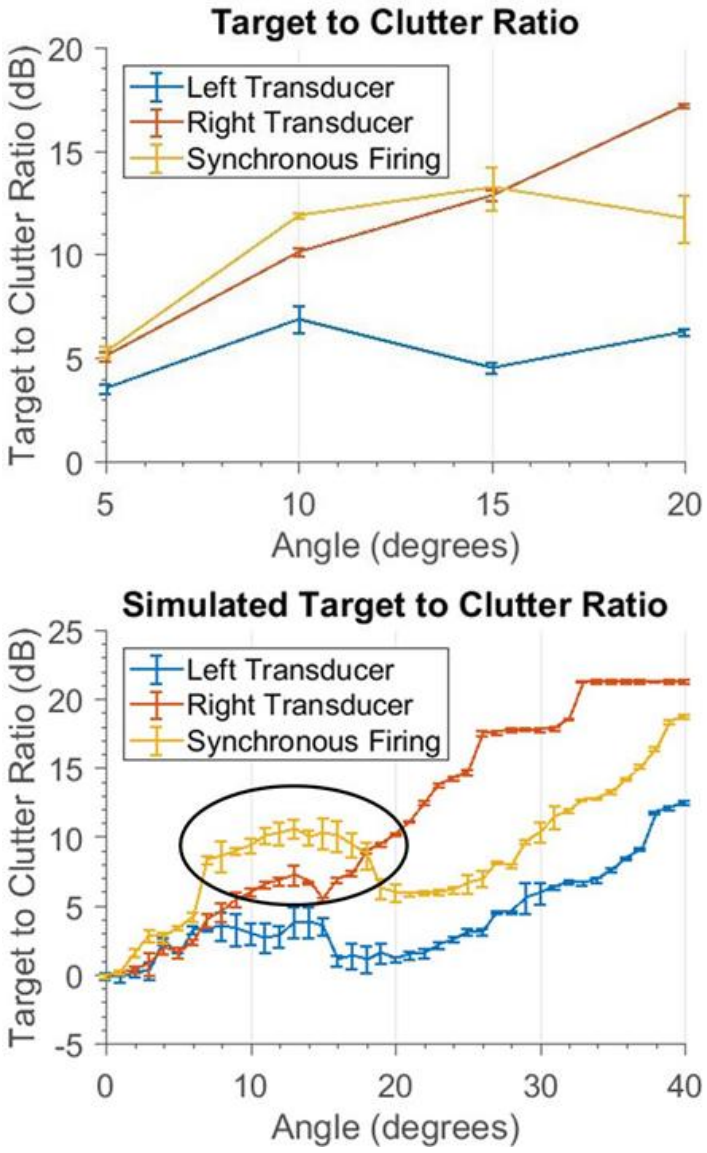

Figure 2.14: These graphs show how a clutter object appears at different angles. The target and clutter objects are at the same range. Only the angle to the clutter object is changed. The top graph shows the ratio of the target and clutter amplitude. The bottom shows simulated data, where only one object was scanned across all of the angles. The ratio was computed using the echo at angle 0 (i.e., the target) and the other angles (i.e., the clutter object). The circled area shows that for angles less than 18 degrees the synchronous firing has better clutter rejection.

2.2.3 Beam Shaping Results

For the beam shaping study, the results are shown in Figure 2.14. The target to clutter (amplitude) ratio is used to normalize the data, which accounts for the difference in magnitude of the different firing patterns. The simulated data was created using the echoes from one real PVC pole recorded across all of the angles. The center measurement is used as the target amplitude; all other angles are treated as clutter amplitudes. The target to clutter ratio is calculated between the center and all other angles.

The synchronous firing pattern has a higher target to clutter ratio than the left or right transducers alone. This only occurs for angles less than 18 degrees. This is due to the side lobes of the interference pattern; once the clutter starts to enter these lobes it is no longer sufficiently rejected, and a single transducer will yield a better target to clutter ratio. In between 6 and 18 degrees, where the most benefit is seen, there is a 3.39dB average increase in the signal to clutter ratio is seen with the synchronous firing pattern compared to the next best single transducer.

2.3 Discussion

2.3.1 Adaptive Delay Discussion

The adaptive delay system for alias rejection tackles a problem that most engineered sonar systems avoid at the cost of a lower sampling frequency. When overlapping echolocation cycles are unavoidable, some form of pulse labeling is most commonly used [8], [16]–[18]. These techniques remove the issue of pulse-echo ambiguity since every pulse has its own unique characteristic. The approaches

presented here are unique in that the pulse-echo ambiguity remains and tracking is maintained in spite of it. This allows a much simpler, single frequency system to be more useful.

The biggest limitation of the adaptive delay system is that it can only deal with a small number of aliased echoes. The case when two aliases sandwich the target is dealt with, but if three or more aliases occur in the right spots, there may be no delay time that prevents the target from being obscured.

2.3.2 Movement Discussion

The movement strategy is much different from the other strategies since it cannot be done on a pulse to pulse basis. Moving the sonar to improve sensing also impacts the decisions of navigation that the sensing is intended to facilitate. These results provide more information to consider by the navigation system that must balance sensing and overall task goals. The basic geometry and the angular response of the sonar system suggest that lateral movement with respect to orientation of the sonar is most effective. Another consideration is that any change in sensing angle may, in fact, generate new aliasing problems as it turns to include new background objects. Note that this approach (like the pulse timing method presented in section 2.1.5) will have little to no effect for clutter objects that appear at the same range as the target.

2.3.3 Beam Shaping Discussion

This technique is a useful way to reduce the effect of aliasing and is the only strategy presented here that is also potentially effective for objects at ranges similar to the target. It is most effective for small angles off-center. Synchronous firing creates a loud central lobe down the central axis of the sonar head. This allows for objects at longer ranges to be detected. This study did not utilize the out of phase firing primarily because the target is assumed to be held in the center of view. The out-of-phase transmission pattern has its minimum in the center of view. If a different tracking algorithm was used that kept the offending clutter in the center, this firing pattern could also be useful in rejecting clutter.

This kind of interference pattern has also been observed to be used by certain bats [15]. Carollia perspicillata emits sound from two nostril holes. These two nostril holes appear to interact in the same way as depicted in the sonar system above.

2.3.4 Combining the Strategies

While these three strategies have been presented and considered separately, they can be combined into an integrated approach. Adaptive delay and beam shaping can be used simultaneously; the delay can be changed independently of the beam shape. Movements to specifically reduce aliasing can also be made, although other factors will likely affect what actions are taken.

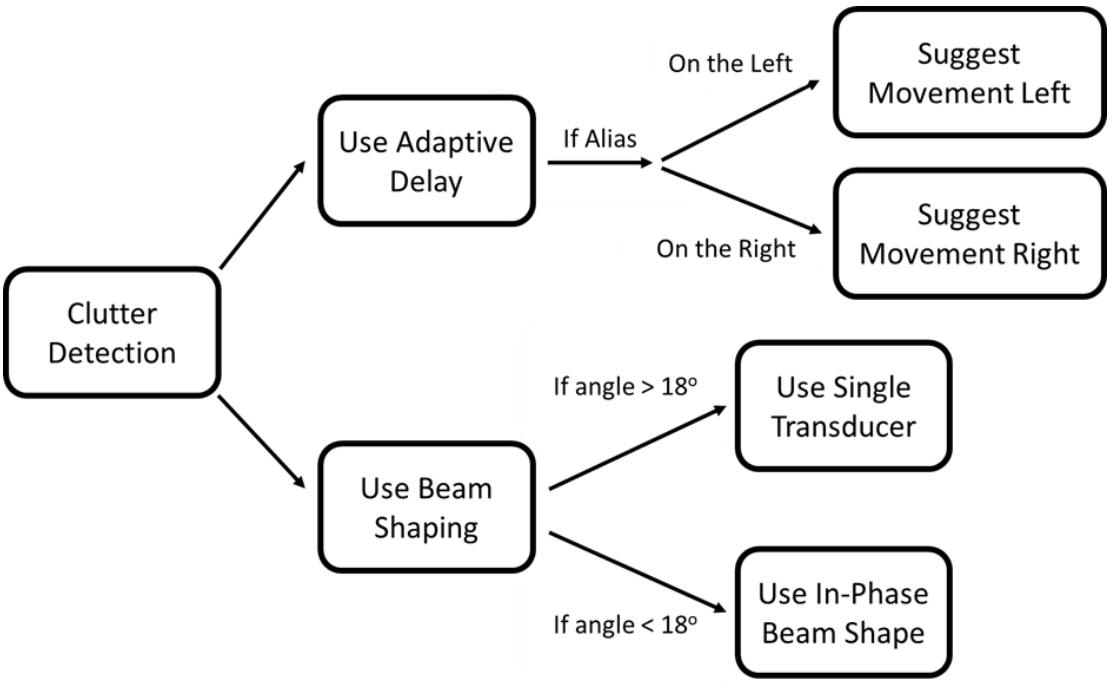

Figure 2.15: Flowchart for integrating the three strategies. Once clutter is detected, beam shaping and the adaptive delay can be used simultaneously. If the adaptive delay determines that the object is an alias, a movement direction will be suggested.

If an alias is detected, the adaptive delay approach can be used to prevent the target from being obscured. At the same time, a movement direction can be suggested based on the apparent angle of the alias. If the obstructing echo is

determined to be a real object and not an alias (part of the adaptive delay code), then different beam shapes can be used depending on the apparent angle of the obstructing echo. If the angle is less than 18 degrees, synchronous firing will be used. If the angle is greater than 18 degrees, only one transducer will be used. This approach is summarized in Figure 2.15.

These strategies complement each other well. Together, they present a multi-pronged approach for dealing with the interference produced while using high pulse

repetition rates. Each strategy is suited to a different situation and need not be used simultaneously.

2.4 Conclusion

Three different active strategies for dealing with echo aliasing are described that can allow the use of sonar at high sampling rates in cluttered environments. Although a time-domain attentional system is assumed to be able to focus on a specific range to track objects, echoes from clutter objects can overlap in time, obscuring or confusing such an attentional system. At very short interpulse intervals, echoes from the background arriving after the next pulse appear to be at a shorter range then they actually are. These "aliases" can overlap the target and interfere, causing a failure of the tracking system. A dynamic pulse-timing strategy is proposed that can effectively "push" or "pull" the aliased echoes away from the tracked target echo by decreasing or increasing the interpulse interval. This prevents aliases from interfering with tracking. We have also presented a method of avoiding or reducing aliases based on positioning, as well as a method of shaping the echolocation beam to reduce the effect of aliasing or clutter.

Bats have been shown to use several different strategies when encountering cluttered situations that require fast sampling. They have been observed to change the frequency content of consecutive pulses [8], alternating between short and long pulses [7], and using the directionality of certain harmonics to focus in a given direction [9]. The system presented here operates on a single carrier frequency, so frequency-based techniques for clutter rejection were not explored, however, we have

shown that other techniques are possible (pulse timing, flight steering, and beam shaping) and are possibly also in use by echolocating bats.

Chapter 3: Echo View Cells from Bio-Inspired Sonar

A neural network model was used to implement echo view recognition that incorporates concepts from machine learning related to pattern separation and classification. A key aspect of this investigation is the attempt to bridge the gap from high-dimensional, low-level, sensory inputs to the more symbolic, discrete nature of place recognition that is critical to higher-level cognitive models of path planning [43]. A key goal is to ensure that the echo view cells respond over a wider area and not just to a single coordinate in space. One limitation of the work is that only limited information is available from the narrowband sonar (typical objects are represented by only a few echoes) and object recognition was difficult, preventing a landmark-based approach, as is common for visual place recognition algorithms. Instead, views were recognized based solely on the spatiotemporal pattern of echoes allowing the memorization of views in a variety of environments without prior training of an object recognition layer. From view recognition, direction-independent place recognition can be constructed in convergence with odometric information. Such approaches to place recognition with sonar have been used [41], [44]. One challenge with sonar is that small changes in the position and angle (particularly in man-made environments) can produce large changes in the resulting echo pattern. Multi-path reflections are also sensitive to positioning. To explore this, data was taken with a large variety of small changes to the positioning of the sonar.

This work explores two very different neural networks that can achieve this: a single layer neural network operating on a recorded echo pattern presented as an image, and a biologically-realistic, spiking neural network (SNN) presented with

echoes in the time domain to simulate live sonar signals. In addition to our motivations to ultimately model and understand the biological implementation of sonar-guided behavior (mentioned above), this work has applications for mobile, autonomous robotics.

3.1 Materials and Methods

3.1.1 Dataset Description

Data was recorded in our laboratory and the adjoining hallway. 66 different recording locations were spread throughout this environment. Locations were spaced 2 feet apart where possible, forming a grid-like placement (Figure 3.1). No attempt was made to restructure the objects in the lab to accommodate the sensing; things were left as they were. No objects were moved during the recording at different locations.

To capture a broader view, a variety of data was collected at different translations and rotations within each square at each of the 66 locations. Across 1 square foot, data was recorded at 25 different translations inside a 5x5 square grid with a 3 inch (7.6 cm) spacing. At each of these 25 points, data was recorded at 11 different angles, ranging from -5 to +5 degrees in 1 degree increments (Figure 3.1B and 3.1C). 10 samples were taken at each angle. In total, each square location has: (25 translations) x (11 angles) x (10 repetitions) = 2750 sonar images per location. With 66 locations, the full data set consists of 181,500 sonar images.

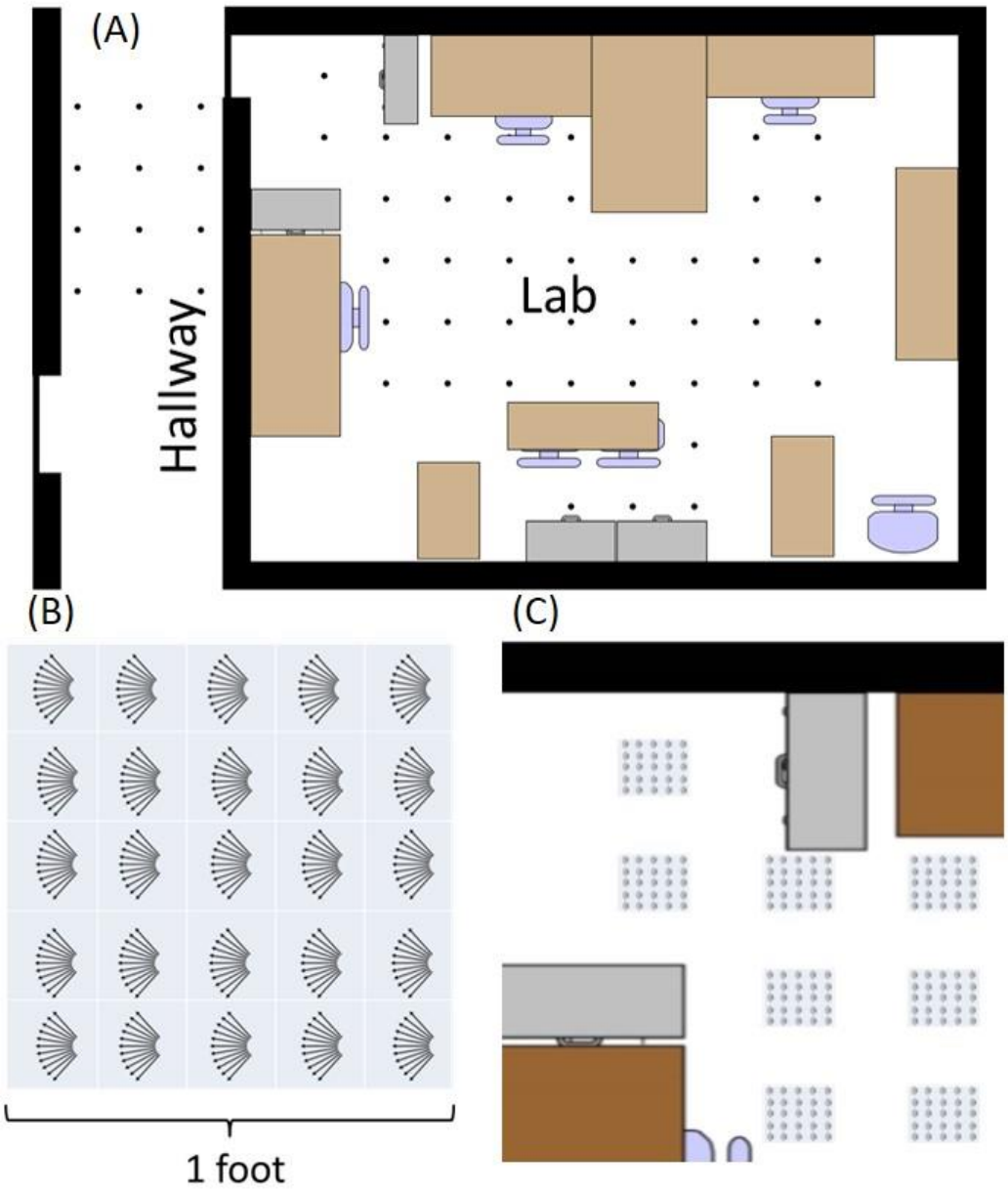

Figure 3.1: The top image (A) shows a map of places data was recorded. Every dot is a recorded place. Locations and objects are approximately to scale. Bottom left (B) shows how a variety of data was recorded at different translations and rotations at each point in the lab. 11 angles were recorded along 5 rows and 5 columns giving 275 recordings at each place. Bottom right (C) shows explicitly how data was recorded at each place, capturing a large variety of data throughout the lab.

3.1.2 Echo Fingerprint Recognition

Two different neural network architectures were tested for their ability to recognize which of the 66 locations a sonar pulse came from. A conventional, single layer network was used and a biologically-plausible, temporally-based architecture called the Synaptic Kernel Inverse Method (SKIM) [45] was used. The inputs and outputs of both networks were similar. The inputs consisted of one sonar image. 255 range bins were used with data from the 3 transducers, resulting in a 765-dimensional input vector. The envelope amplitude data was supplied to the network. If there was no echo in a time bin, the value was kept as zero. The resolution of each range bin was 2.14cm or 0.84in. Each sonar image was L2 normalized before being fed to the network. While normalizing means the network does not have direct access to the echo magnitudes, the relative magnitude between echoes contains more reliable and reproducible information, such as the magnitude difference between transducers which relates to echo direction. Each output corresponds to a different location, so with 66 locations there are 66 outputs. In both networks, a form of supervised learning was used to train the network.

Although the angle of an arriving echo could be calculated using the magnitude difference between the transducers (e.g., using interaural level differences) to reduce the dimensionality, we chose to retain the raw values and let the network learning rules determine how this information would be used.

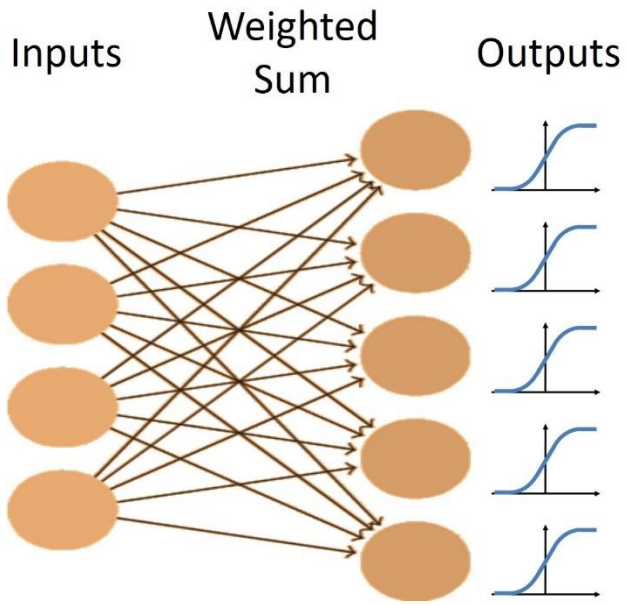

Figure 3.2: The network architecture for the single layer network. There is one layer of fully connected weights from the inputs to the outputs. Each output has a logistic nonlinearity applied to it to maintain outputs between 0 and 1.

3.1.3 Single Layer Feedforward Network

In this experiment, a very simple neural network was used to process the data. The network consisted of the input layer fully connected by weights to the output layer (Figure 3.2). The nonlinear logistic function was applied to the summation of weighted inputs to provide the output. Learning was performed by a modified version of gradient descent that uses an adaptive momentum term to speed learning, called the AdamOptimizer algorithm [46]. This was implemented in the machine-learning software package, TensorFlow [47] on Google's Colaboratory cloud computing platform [48], allowing us to speed up the training with free use of their GPUs.

In this task, the single layer network performed as effectively as multiple-layer networks and its simplicity led to an easier observation and analysis of how the network was solving this problem.

3.1.4 Synaptic Kernel Inverse Method (SKIM)

SKIM is a multi-layer network architecture that combines the benefits of Extreme Learning Machines (ELM) but with spiking neuron (temporal) representations. Sonar lends itself to being represented in the spiking domain because echoes themselves are inherently time-based signals and typically pulsatile in nature. The temporal nature of this network suggests a real-time implementation using spiking neuromorphic hardware. Figure 3.3 illustrates the SKIM network architecture [45].

The first layer of weights in the SKIM network consists of fixed, random weights connecting the inputs to the hidden layer. These weights can be positive or negative. The fanout here is usually 10-20 (or a hidden layer that has 10-20 times more neurons than the input layer), resulting in a very large hidden layer. This is typical of an ELM approach, which aims to expand the dimensionality of the input data to make pattern separation easier [49]. There is also a nonlinearity applied at each hidden unit. Every hidden unit has a randomly selected temporal synaptic kernel associated with it that consists of a time delayed alpha function. If A is the activation of the unit, t is the time, ΔT is the delay, and τ the width of the alpha function, the equation is:

$$\boldsymbol{f(t) = tanh\,(A\,\frac{t-\Delta T}{\tau}e^{-\frac{t-\Delta T}{\tau}})}$$

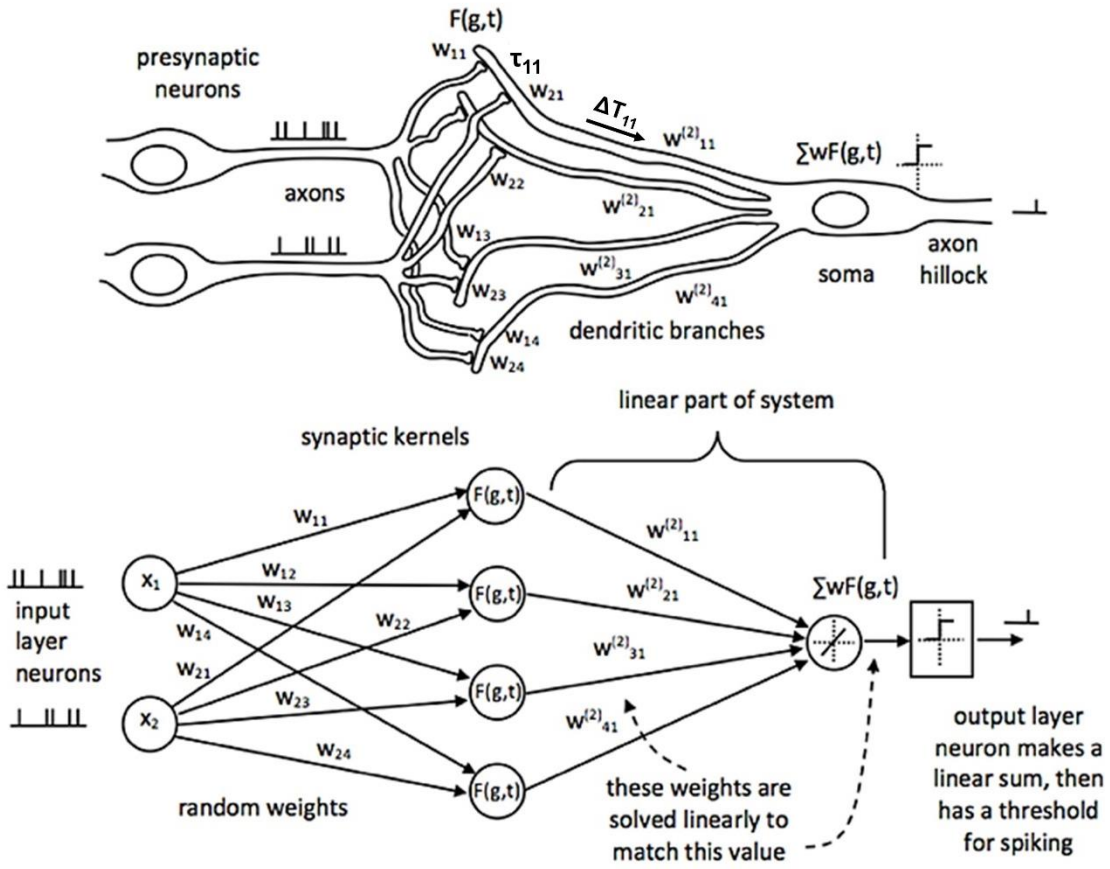

Figure 3.3: Adapted from [45]. The architecture for the SKIM neural network. The top of the figure shows what a corresponding biological system would look like, while the bottom shows this network from a computational perspective. Inputs from the presynaptic neurons are summed onto the dendrites of the postsynaptic neurons. Each dendrite has an associated nonlinear, synaptic kernel (F(g,t)) with a time constant (τ), and dendritic delay (ΔT). The dendritic activity is summed onto the soma and creates a spiking output when above a threshold. The weights from the input layer to the hidden layer are static (wxy); the linear connection from the hidden layer to the output has weights that are trained (w(2)yz).

where different hidden units have different delays (ΔT) and widths of the alpha function (τ) (Figure 3.4). The time delay is essential to recognizing patterns that occur over time, and gives the network a form of memory, a way to be influenced by data in the past. A compressive nonlinearity (the hyperbolic tangent, tanh) is applied as well.

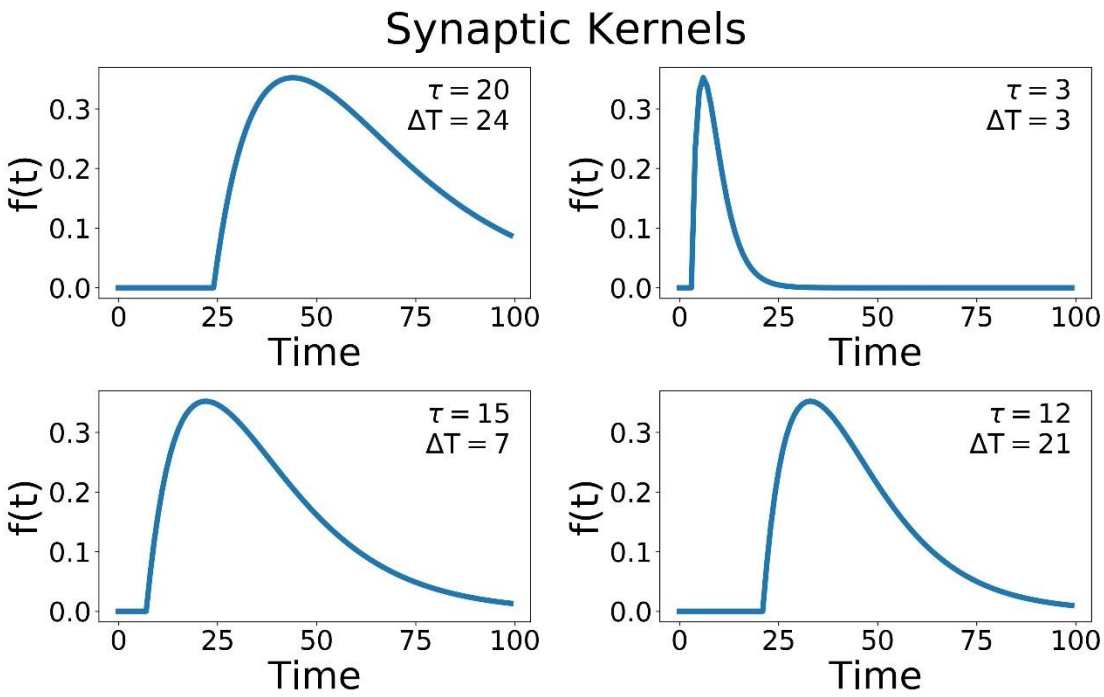

Figure 3.4: Some example synaptic kernels. Two parameters are changed, the delay for the onset of the function (ΔT), and the width of the alpha function (τ). The x-axis corresponds to the variable t of this function.

These hidden units create a high-dimensional, nonlinear transformation of the input data that has occurred recently in time. This allows for complex, temporal patterns to be more easily recognized and separated.

The next layer of this network is linear. There are a set of fully-connected weights from the hidden layer units to the output. These are the weights that are modified during learning. Since this is the only dynamic part of the network, the learning is simplified. As this is a linear transformation with a known hidden-unit activation and a known output (since we are performing supervised learning), the weights can be solved for analytically.

If M is the number of hidden units and k is the number of time steps in our dataset, we obtain a matrix describing the hidden unit activation over time, $H \in \mathbb{R}^{Mxk}$.

If N is the number of outputs, we have the output activation matrix, $Y \in \mathbb{R}^{Nxk}$. The weights connecting the two layers will be $W \in \mathbb{R}^{NxM}$, such that $WA=Y$. To find the weights we simply have to solve for W, giving $W=YA^{+}$, where A^{+} can be found by taking the Moore-Penrose pseudoinverse of A.

To solve this analytically, we use the Online PseudoInverse Update Method (OPIUM) [50]. This is an application of Greville's method, which shows an incremental solution to finding the pseudoinverse, but is adapted and simplified for this specific problem to reduce the needed computation without losing accuracy.

3.2 Results

3.2.1 Single Layer Network

This network was trained to predict which of the 66 locations a sonar pattern came from. The recorded sonar dataset was split into three parts, 80% training data, 10% testing data, and 10% validation data. The data was randomly shuffled across locations and positions within locations before being split into these three groups. Our accuracy of identifying the location of a particular pattern from the validation data set reached 97.5%. A graph of the accuracy across the training regimen is shown in Figure 3.5.

Since this network is very simple, it is easy to understand how the weights can be interpreted. Each output neuron has a weight corresponding to every input. These can be thought of as the perceptive field of this output neuron. By looking at which inputs cause the output to activate, we can get an idea of the sonar image preferred by each output neuron. Figure 3.6 shows some example weights from the network. One

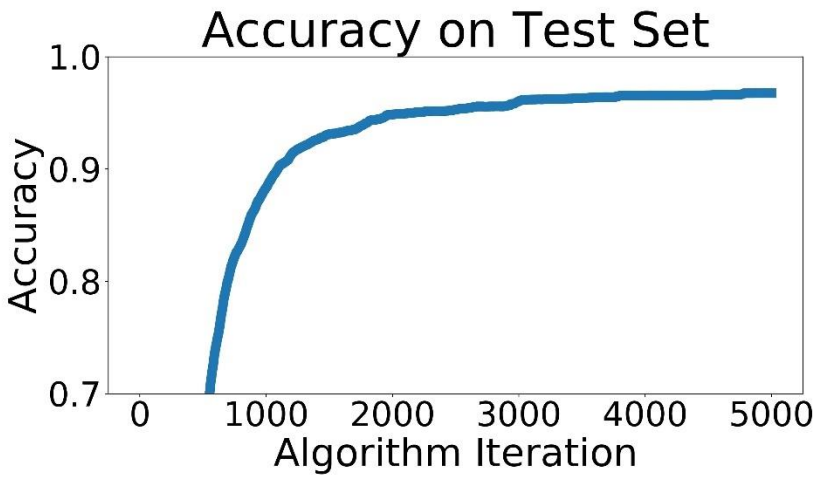

Figure 3.5: Network accuracy as training progresses. Each algorithm iteration takes approximately 0.5 seconds, and the network takes about 1 hour to train.

noticeable pattern in these weights is the splitting that occurs between the right and left transducers; there are clear ranges where one will be positive and the other will be negative. Functionally, this is the network learning to look for objects at a certain angular orientation. Another clear pattern that arose in the network weights; the weights from the hallway seemed to be synchronized across transducers (Figure 3.7). These weights were also lower in amplitude than those from inside the lab.

Figure 3.8 shows how the different view cells responded across the whole map. It is clear that the network learned very rigid boundaries where it was trained to do so. Although this demonstrates a successfully trained network, the sharp distinctions between neighboring locations is not what is seen in mammalian place cells.

3.2.2 SKIM

In the SKIM network trained with OPIUM, we achieved up to 93.5% accuracy on our dataset. The choice of time constants (τ, the alpha function widths) and delays

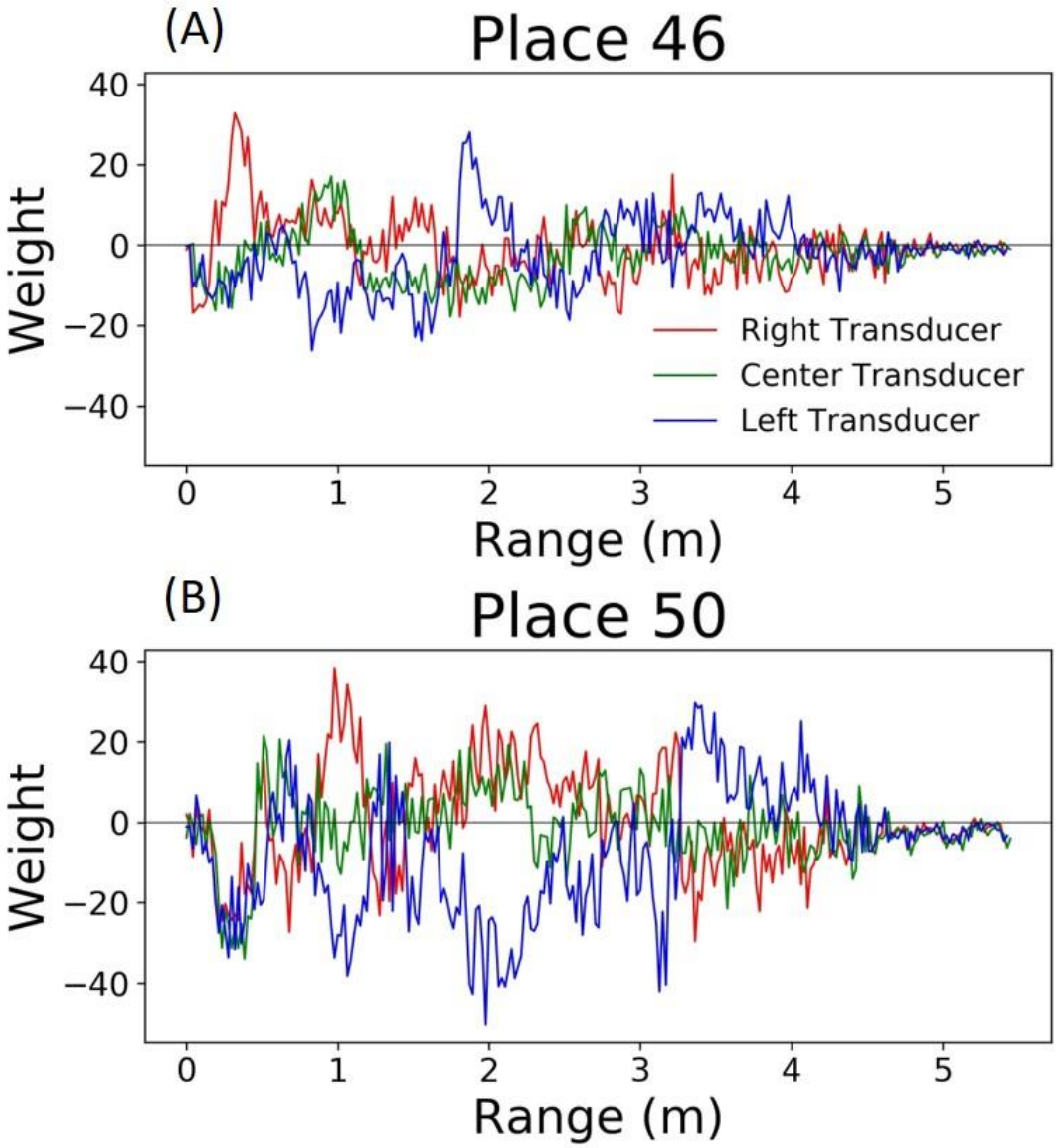

Figure 3.6: Perceptive fields of the output neurons in the single layer network. It's clear that in some spots these perceptive fields split the left and right signals. This gives the network the ability to discriminate direction.

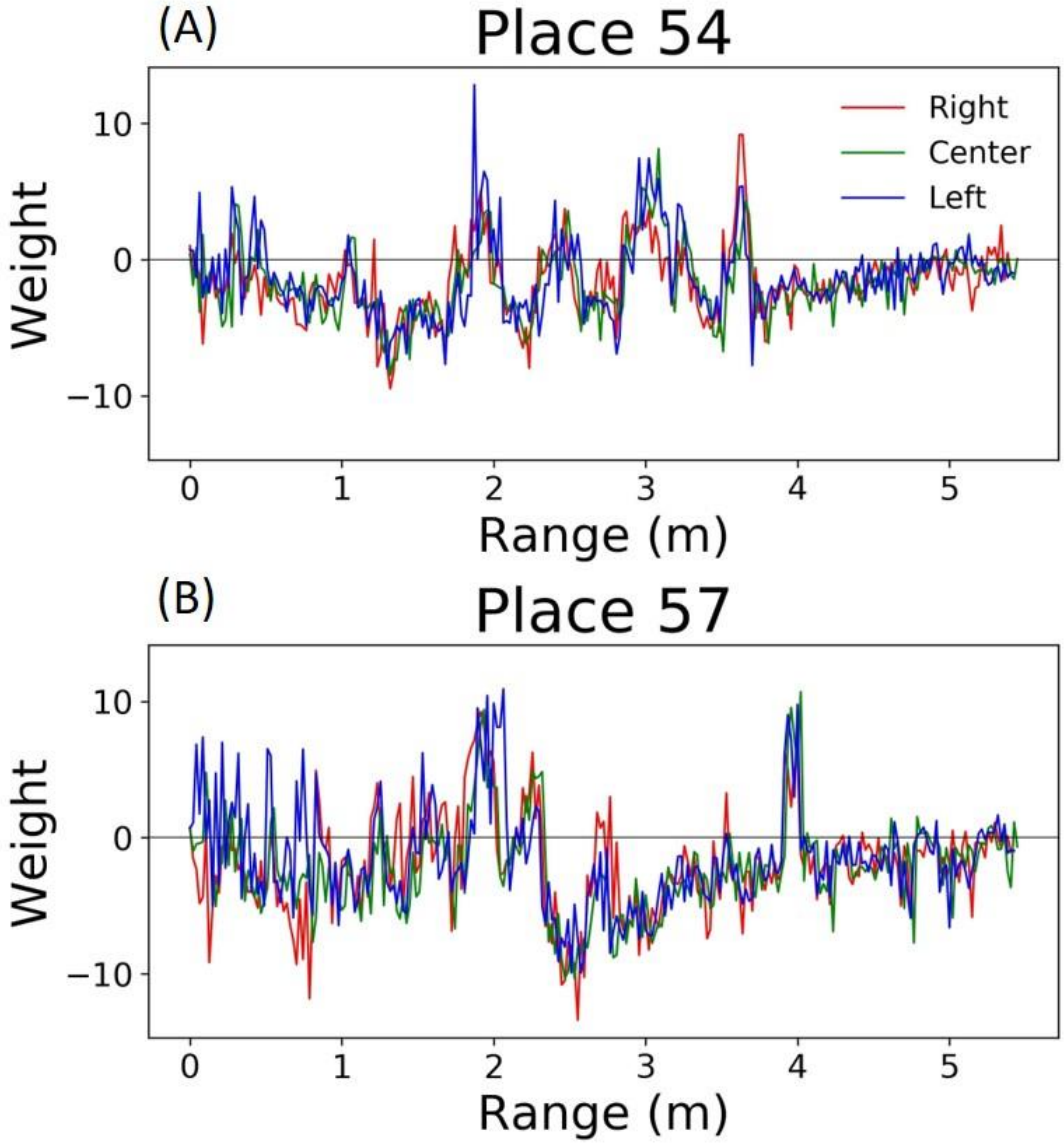

Figure 3.7: Perceptive fields of the output neurons in the single layer network. These are from the hallway data. These weights were of lower amplitude, and all transducers were correlated with one another.

(ΔT) for the synaptic kernels was very important. The time constants determine the temporal precision the network can observe; large time constants lead to less temporal precision. Long time constants provide tolerance to temporal jitter between patterns but result in a loss of temporal discrimination when needed. The time constants used for this network covered one to five time bins, with τ's randomly chosen between 0.5 and 1.5, keeping a relatively narrow and precise response. The choice of delays determined which temporal part of the data is relevant (i.e., beginning, middle, end of the pulse). The delays were distributed randomly over the length of the sonar pulse to ensure that all the echoes had an equal probability of activating the network, with ΔT's randomly chosen between 0 and 255. The network was trained to deliver an output at the end of a sonar image (t=255). Accuracy was determined by taking the output neuron with the highest activation at t=255. Figure 3.8I-3.8L shows how the SKIM view cells responded across the whole map. The response is very similar to the single layer network with rigid boundaries between views.

3.2.3 Recognition Outside of Training Data

Outside of the locations (squares) where data was collected, both networks does not predictably recognize that it is near a known location. The accuracy was high when in an area it was trained on, but recognition drops quickly even inches away. Figure 3.8 shows this for the single layer network; Figure 3.9 shows this for the SKIM network. To spread the activation of the network to neighboring areas outside the training area, network training was changed. Instead of an output neuron being trained to 1.0 in its corresponding location and all other neurons trained to 0.0,

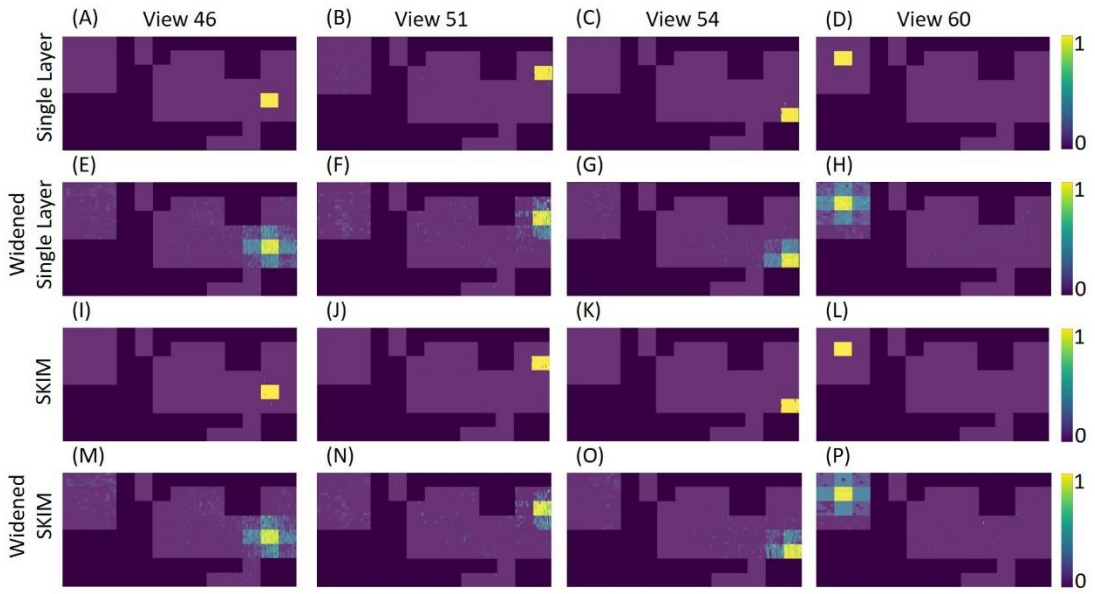

Figure 3.8: An overhead view of the different echo view fields created by the two networks. This map is the same as shown in the top of Figure 3.1. Each plot represents a different echo view cell's activation across the entire map. The top plots (A-D) show the original network for four different views, and the plots (E-H) shows the single layer network with widened labels, resulting in neighboring views being activated. Plots (I-P) show the view activations for the SKIM network and the widened SKIM network. It is important to note that only areas in the training dataset are displayed. The 1 foot squares in between each of the locations have been omitted (shown explicitly in Figure 3.1C).

neighboring neurons were trained to respond to neighboring views. A Gaussian function was used, giving adjacent views an activation of .5 and diagonal views and activation of .38. After this round of training the accuracy of the single layer network dropped to 92.3%, while the accuracy of the SKIM network remained stable at 93.4%. Figure 3.9D-3.9F and 3.9J-3.9L show the results of this new training for the single layer network and SKIM network respectively. The new activation pattern of the network is now spread through areas that were not explicitly trained on, and

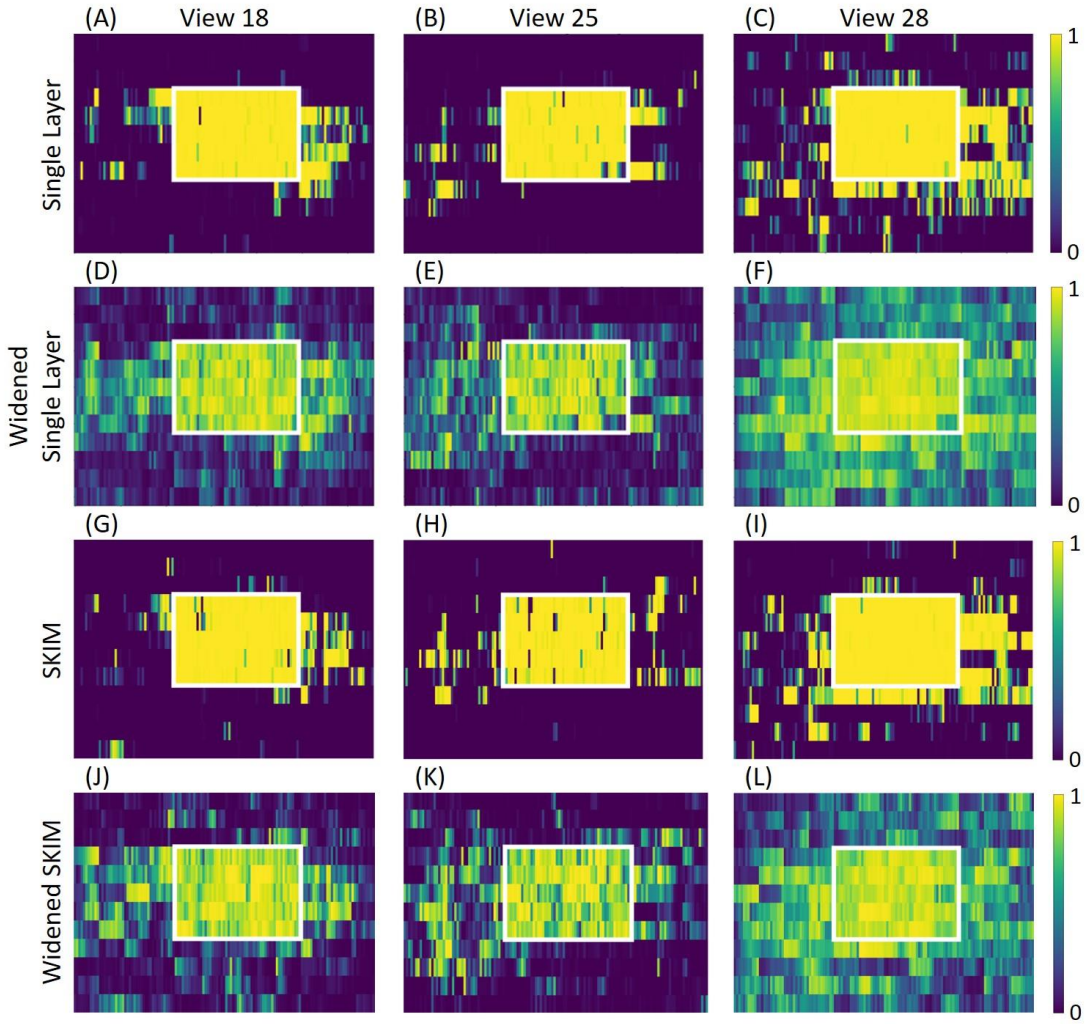

Figure 3.9: Each graph presents an overhead view of a location. Inside the white square is where training data was recorded; outside the white square is an adjacent area that was not used for the training of the networks. Along the x axis, eleven adjacent pixels show the eleven angles for each of the 25 (5x5) spatial positions inside the white box. Pixels along the y axis are spaced evenly. These view neuron activation patterns are generated by the corresponding output neuron from the neural network. The top plots (A-C) show how the single layer network responds around these locations, showing sparse activation outside the trained square and very high activation inside the square. Plots (D-F) show the single layer network trained to respond to neighboring views. Plots (G-L) show the same information, but for the

Figure 3.9 (continued): SKIM network and the widened SKIM network. The widened networks show a much more spread out activation in the non-trained area outside the square.

qualitatively looked more like biological place fields. Figure 3.8E-3.8H and 3.8M-3.8P also shows how these new view cells respond across the whole map. There is now more noticeable activation in areas that were not trained on. The cells have become much more broadly tuned. We call this new network the 'widened' network, in contrast to the 'original' network. The single layer network and the SKIM network responded very similarly in all the cases presented.

3.3 Discussion

3.3.1 Functionality Test Along a Path

To demonstrate how this system might be used in practice, sonar data was recorded along a path consisting of points both inside and outside of the training data. The single layer network's response to this data shows how views can be recognized along the entirety of this path (Figure 3.10).

The widened network, which allows multiple view neurons to be active at once, creates a broader, more spatially-continuous response when compared with the original network. Less reliance on a single view neuron activating provides a more stable and nuanced interpretation of location. In situations where the original network fails to activate the correct view neuron, the widened network is more likely to alleviate the situation by activation of other nearby view neurons.

Leaky integration was also used to help smooth out the network response over time; each activation is given an exponentially decreasing tail over time. With A_t as

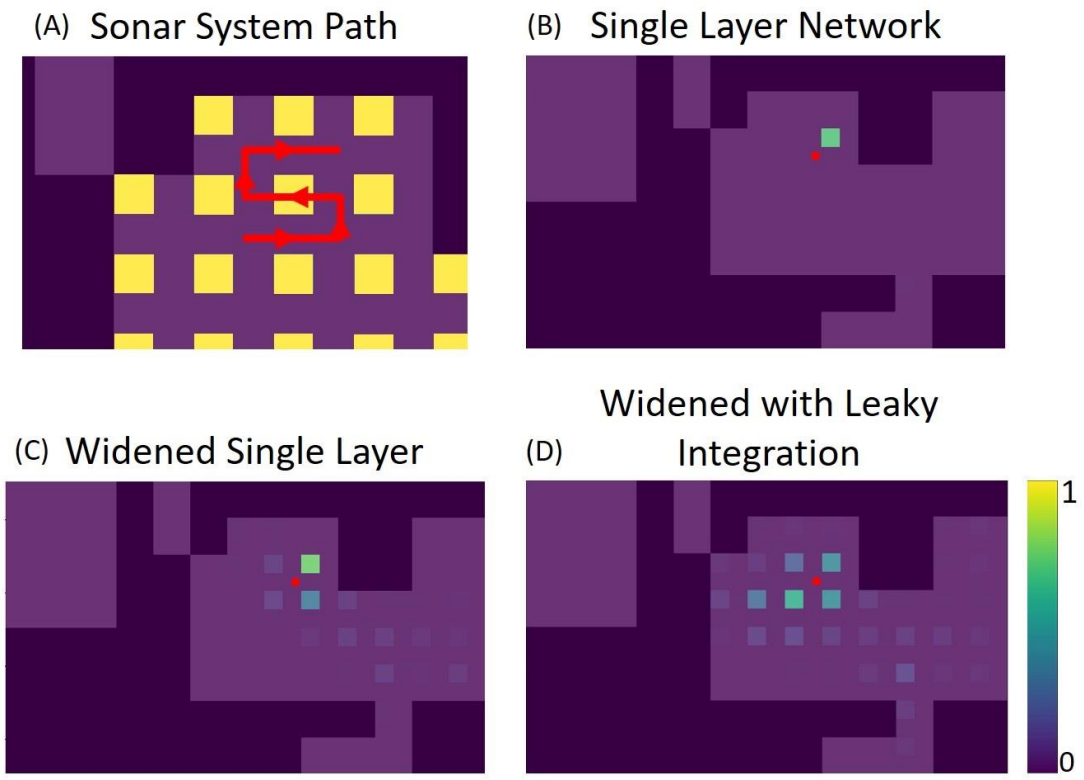

Figure 3.10: Panel (A) shows the path the sonar system moves through, in red. There are 39 positions total along this path, each position 3 inches from the last. The portion of the path within the yellow squares is contained in the training set for the networks (5 of the path positions). The rest of the path was not used for the training of the networks. Panels (B-D) show echo view field responses on the path. The red dot represents the position of the sonar. The activations of the echo view cells are shown in their corresponding location, seen as colored squares on the plots. The single layer network was trained to have only one view cell active at a time. The widened network allows for more cells to be active at once, improving accuracy in between trained views. The leaky integration maintains a more stable activation due to its use of the past activations in the path.

the activation for a position at time t, and L_t as the activation for a position after leaky integration is applied, the equation used is $\boldsymbol{L_t = \alpha A_t + (1 - \alpha) L_{t-1}}$. In this example, one view is about 10 movements wide. Using a leaky integration constant ($\boldsymbol{\alpha}$) of 5/9

allows for activation to be maintained at %10 of its original value 10 time steps in the future, allowing persistent activation while moving across a position at the cost of a slight lag. An equivalent way to calculate this would be to have each activation exponentially decaying over time; the corresponding time constant would be 4.5. In some locations on the path, the sonar is not able to correctly recognize the view. For this example, integration over time gives the network more stability and accuracy.

The widened network with leaky integration gives consistently accurate results over the whole path. The echo view fields activated are generally smooth over space and decaying activation can be seen multiple locations away. To evaluate the effectiveness of these echo view fields, we calculated the activity-weighted centroid at each point on the path, giving us an average point of each field to compare with the actual position of the sonar. The distance between the activity-weighted centroid and the actual position was used to calculate a mean error. Across 117 steps along 3 different paths, the original network's average error was 28.6 inches (72.6 cm), the widened network's average error was 18.6 inches (47.2 cm), and the widened network with leaky integration's average error was 16.3 inches (41.4 cm). This system successfully recognized locations that are not contained in the training set; the network can generalize and recognize many nearby views. When this fails, leaky integration allows past information to maintain a stable sense of place for the system.

Supplementary videos show the activations of the original network, the widened network, and the leaky integration applied to the widened network

similar to Figure 3.10, but over the entire path, and can be viewed at https://www.frontiersin.org/articles/10.3389/fnbot.2020.567991/full#supplementary-material.

3.3.2 Context/Previous Studies

These results complement previous studies that have used sonar to aid in place recognition. A large inspiration for our project was BatSLAM [40], a biomimetic sonar system that used odometry and sonar to map an area of their laboratory. Because odometry is quite inaccurate due to wheel slippage and other errors, such as compounding inaccuracies in estimating direction and position, sonar was used to provide error correction. Their system first drew paths of motion based solely on odometry. When the sonar-based recognition system recognized the current location from a prior visit, it updated the odometry system to match its memory and propagated the correction to earlier time steps for consistency. This was sufficient to correctly create a map of the area with little error. While this approach showed that sonar was able to aid place recognition, it did not do so in a biologically-plausible manner. Over the robot's path, 6000 sonar measurements were taken, and 3300 different places were established. While this system provides a method to maintain an estimate of the robot's position, it does not seem to reflect what little is known about how biological memories of the environment. Memorizing 3300 different places all within one environment is computationally and memory-intensive; it is not a biologically-plausible algorithm. While our study attempted to show that odometry is not needed for view recognition, incorporating odometric information can provide a strong framework for unsupervised mapping. For example, a new 'place' can be

created when a system, using odometry, estimates it is a certain distance from any other ‘place’.

Another recent paper explored the idea of recognizing place with sonar in three different locations [41]. Using a very precise sonar sensor they measured the echo response at positions over a wide range of angles and along a linear, 10m long path. They collected an enormous amount of data (over 20,000 echo traces) and evaluated whether the echoes varied smoothly over angle and distance as well as whether unique locations could be classified. Most of the data came from angular variation; large translational steps contrast the high angular resolution. They also found places that were difficult to distinguish between, mainly in open areas with few objects to sense, but concluded that sonar is enough to recognize most locations. When they were comparing different positions along a linear path, they compared the same precise angle (0.1 degree error) from the different positions. This is much more precise than an animal can hope to achieve, in reality both angle and position will be changing at the same time. We have shown in this study how sensitive an echo signature can be to changes in angle; we expect place recognition to be tolerant to moderate changes in the sensing direction. Our study can complement this one by providing a wider, two-dimensional range of positions for comparison as well as removing the need for very precise angular measurements.

In our study, all views were looking in the same direction. A network that could respond to views in different directions but at the same general location would be a step towards modeling a more general place cell. This could be modelled using an additional layer of a neural network. We have shown that different views can be

separately recognized in a single layer network, another layer would be able to select which views correspond to the same place. This could be as simple as an 'or' function that allows a view from any direction to activate the place cell.

3.3.3 Single Frequency vs Broadband

One important aspect of the sonar currently used in our system that is not biologically-realistic is the use of a single frequency (40 kHz). Bats use a broadband sonar pulse that provides much richer echo signatures with spectral content that likely contributes to object characterization that is not possible with our sonar [51]. Even with this limitation, this study shows that place field generation is still possible knowing only object range (inferred by the peak sound pressure on the three transducer channels) and echo magnitude. Different objects with multiple close surfaces can also produce echoes with different durations. With a broadband sonar sensor, it may be possible to significantly improve the size and reliability of the place fields.

3.4 Conclusion

We have presented a robotic sonar system that uses ultrasonic transducers to mimic bat echolocation and have demonstrated two different networks that can recognize sonar views over a range of angles and offsets ('echo view fields'), with one network showing that this can be done in a biologically plausible manner. This view-based approach does not require the identification of specific objects or explicit use of landmarks. The echo view cells produce "reasonable" responses outside of

places where training data was collected and has the potential to be integrated into a larger system to model bat hippocampal place cells and spatial mapping.

Chapter 4: Conclusion

This work has demonstrated the use of a bio-inspired sonar system for two different tasks: tracking amid clutter and view-based place recognition. These tasks correspond with actions and behaviors that bats need to use in the wild on a daily basis. Tracking is an essential tool for hunting and prey capture; place recognition is an essential part of being able to navigate and explore an environment. These studies help us to understand parts of these more complex behaviors and give us a small hint of what it may be like to act based on echolocation rather than vision. Still, these controlled and isolated experiments are unable to capture the complexity of an actual bat's life, where these skills and many more need to be used in concert, where any actions taken affect many aspects of the animal's life.

4.1 Defining a Place

One of the limitations of the study on view-based place recognition is that places were chosen very rigidly without regards to the data taken. The one foot squares used have little relation to how an animal may choose to define place. From our day to day experience as humans, some places do have rigid boundaries (rooms separated by walls, properties separated by fences or streets), but there are other natural ways to designate a place. While the view based approach doesn't represent individual objects, landmarks are a common navigational tool. Recognizing a specific object can give an identifier for a specific place. In a more statistical approach, features of a view can also be used to identify place. These features may be related to a single object, but also can be more general, such as a cluttered area vs

an open area, a large room vs a small room, a prevalence of softer objects or harder objects, louder echoes or softer echoes.

We have looked for features in our sonar data without much luck. Time was spent trying to find patterns within the spacing between echoes that yielded little results. Originally when analyzing the data, a two layer neural network was used, hoping that the middle layer would act as a feature detector. There were no interpretable features found in this layer, and the network was shown to work just as well without this extra layer. Much like in [3], an alternate way to look for features would be in individual echoes. Looking at the FFT of an individual echo provides much richer information that may be more easily categorized. Since our system is a single frequency and only looks at the peak magnitude, this was not our focus. Attempts at object recognition using the features of an individual echo is an interesting topic worthy of further research.

Another way to define a place is by the tasks that can be accomplished there. People define many places in this manner, a bedroom for sleeping, an office for working, a kitchen for food preparation. This method of place identification involves more than a view or snapshot of a place (which this study was restricted to). Another line of future study could be to use mock tasks to define a place and then use the sensory experience throughout this location to recognize the place. This would also change the learning methods from supervised to unsupervised, letting the environment and task dictate the boundaries of a place in a more natural way instead of the researcher. One of the tasks that may be used in this manner is navigation. Places can be defined by navigational tasks such as choosing a direction which sets a

crossroads as a place, or even traveling forwards which gives paths or hallways as places.

Another way to define place would be using odometry. This study did not integrate odometry with sonar sensing, but in general odometry is an important tool for any navigational task. In the study along a path, we used integration over time to give a smoother, more continuous activation of place. It would be just as appropriate to use an odometric system to track how quickly position is changing and adjust the importance of nearby activations. This could also be used to create an unsupervised learning method, where setting a distance between different places would allow the creation of new places.

4.2 Conclusion

We have shown that places are able to be distinguished from one another using only sonar. This demonstrates that information gained from sonar is rich enough to inform decisions about navigation and other tasks without the aid of other vision or odometric sensors. Exploring how to use sonar and how bats use echolocation will provide useful tools for guiding autonomous vehicles, scanning new areas, and object tracking.

Bibliography

[1] J. O. Whitaker and S. L. Gummer, “Hibernation of the big brown bat, Eptesicus fuscus, in buildings,” *J. Mammal.*, vol. 73, no. 2, 1992, doi: 10.2307/1382062.

[2] N. Jeong, H. Hwang, and E. T. Matson, “Evaluation of low-cost LiDAR sensor for application in indoor UAV navigation,” in *2018 IEEE Sensors Applications Symposium, SAS 2018 - Proceedings*, 2018, vol. 2018-January, doi: 10.1109/SAS.2018.8336719.

[3] I. Eliakim, Z. Cohen, G. Kosa, and Y. Yovel, “A fully autonomous terrestrial bat-like acoustic robot,” *PLoS Comput. Biol.*, vol. 14, no. 9, 2018, doi: 10.1371/journal.pcbi.1006406.

[4] J. J. Leonard and H. F. Durrant-Whyte, “Application of multi-target tracking to sonar-based mobile robot navigation,” in *Proceedings of the IEEE Conference on Decision and Control*, 1990, vol. 6, doi: 10.1109/cdc.1990.203365.

[5] Y. Petillot, I. T. Ruiz, and D. M. Lane, “Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar,” *IEEE J. Ocean. Eng.*, vol. 26, no. 2, 2001, doi: 10.1109/48.922790.

[6] W. M. Masters, S. C. Jacobs, and J. A. Simmons, “The structure of echolocation sounds used by the big brown bat Eptesicus fuscus: Some consequences for echo processing,” *J. Acoust. Soc. Am.*, vol. 89, no. 3, 1991, doi: 10.1121/1.400660.

[7] A. E. Petrites, O. S. Eng, D. S. Mowlds, J. A. Simmons, and C. M. Delong, “Interpulse interval modulation by echolocating big brown bats (Eptesicus

fuscus) in different densities of obstacle clutter," *J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol.*, vol. 195, no. 6, 2009, doi: 10.1007/s00359-009-0435-6.

[8] S. Hiryua, M. E. Bates, J. A. Simmons, and H. Riquimaroux, "FM echolocating bats shift frequencies to avoid broadcast-echo ambiguity in clutter," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 15, 2010, doi: 10.1073/pnas.1000429107.

[9] M. E. Bates, J. A. Simmons, and T. V. Zorikov, "Bats use echo harmonic structure to distinguish their targets from background clutter," *Science (80-.).*, vol. 333, no. 6042, 2011, doi: 10.1126/science.1202065.

[10] D. M. Wisniewska, M. Johnson, P. E. Nachtigall, and P. T. Madsen, "Buzzing during biosonar-based interception of prey in the delphinids tursiops truncatus and pseudorca crassidens," *J. Exp. Biol.*, vol. 217, no. 24, 2014, doi: 10.1242/jeb.113415.

[11] J. Parker *et al.*, "Genome-wide signatures of convergent evolution in echolocating mammals," *Nature*, vol. 502, no. 7470, 2013, doi: 10.1038/nature12511.

[12] C. F. Moss and A. Surlykke, "Auditory scene analysis by echolocation in bats," *J. Acoust. Soc. Am.*, vol. 110, no. 4, 2001, doi: 10.1121/1.1398051.

[13] C. F. Moss, C. Chiu, and A. Surlykke, "Adaptive vocal behavior drives perception by echolocation in bats," *Current Opinion in Neurobiology*, vol. 21, no. 4. 2011, doi: 10.1016/j.conb.2011.05.028.

[14] J. A. Simmons, "Temporal binding of neural responses for focused attention in

biosonar," *Journal of Experimental Biology*, vol. 217, no. 16. 2014, doi: 10.1242/jeb.104380.

[15] D. J. Hartley and R. A. Suthers, "The sound emission pattern and the acoustical role of the noseleaf in the echolocating bat, Carollia perspicillata," *J. Acoust. Soc. Am.*, vol. 82, no. 6, 1987, doi: 10.1121/1.395684.

[16] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor - System description, issues and solutions," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2004, vol. 2004-January, no. January, doi: 10.1109/CVPR.2004.291.

[17] M. I. Skolnik, *Radar Handbook*. 2008.

[18] N. Matsuta, S. Hiryu, E. Fujioka, Y. Yamada, H. Riquimaroux, and Y. Watanabe, "Adaptive beam-width control of echolocation sounds by cf-fm bats, rhinolophus ferrumequinum nippon, during prey-capture flight," *J. Exp. Biol.*, vol. 216, no. 7, 2013, doi: 10.1242/jeb.081398.

[19] J. O'Keefe, "Place units in the hippocampus of the freely moving rat," *Exp. Neurol.*, vol. 51, no. 1, 1976, doi: 10.1016/0014-4886(76)90055-8.

[20] M. Gil *et al.*, "Impaired path integration in mice with disrupted grid cell firing," *Nature Neuroscience*, vol. 21, no. 1. 2018, doi: 10.1038/s41593-017-0039-3.

[21] J. O'Keefe, N. Burgess, J. G. Donnett, K. J. Jeffery, and E. A. Maguire, "Place cells, navigational accuracy, and the human hippocampus," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 353, no. 1373, 1998, doi: 10.1098/rstb.1998.0287.

[22] E. T. Rolls, "Spatial view cells and the representation of place in the primate

hippocampus," *Hippocampus*, vol. 9, no. 4. 1999, doi: 10.1002/(SICI)1098-1063(1999)9:4<467::AID-HIPO13>3.0.CO;2-F.

[23] N. Ulanovsky and C. F. Moss, "Hippocampal cellular and network activity in freely moving echolocating bats," *Nat. Neurosci.*, vol. 10, no. 2, 2007, doi: 10.1038/nn1829.

[24] M. M. Yartsev and N. Ulanovsky, "Representation of three-dimensional space in the hippocampus of flying bats," *Science (80-.).*, vol. 340, no. 6130, 2013, doi: 10.1126/science.1235338.

[25] M. Geva-Sagiv, L. Las, Y. Yovel, and N. Ulanovsky, "Spatial cognition in bats and rats: From sensory acquisition to multiscale maps and navigation," *Nature Reviews Neuroscience*, vol. 16, no. 2. 2015, doi: 10.1038/nrn3888.

[26] M. J. Wohlgemuth, J. Luo, and C. F. Moss, "Three-dimensional auditory localization in the echolocating bat," *Current Opinion in Neurobiology*, vol. 41. 2016, doi: 10.1016/j.conb.2016.08.002.

[27] K. Diba and G. Buzsáki, "Forward and reverse hippocampal place-cell sequences during ripples," *Nat. Neurosci.*, vol. 10, no. 10, 2007, doi: 10.1038/nn1961.

[28] G. Dragoi and S. Tonegawa, "Preplay of future place cell sequences by hippocampal cellular assemblies," *Nature*, vol. 469, no. 7330, 2011, doi: 10.1038/nature09633.

[29] T. Solstad, C. N. Boccara, E. Kropff, M. B. Moser, and E. I. Moser, "Representation of geometric borders in the entorhinal cortex," *Science (80-.).*, vol. 322, no. 5909, 2008, doi: 10.1126/science.1166466.

[30] F. Savelli, D. Yoganarasimha, and J. J. Knierim, “Influence of boundary removal on the spatial representations of the medial entorhinal cortex,” *Hippocampus*, vol. 18, no. 12, 2008, doi: 10.1002/hipo.20511.

[31] J. S. Taube, R. U. Muller, and J. B. Ranck, “Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis,” *J. Neurosci.*, vol. 10, no. 2, 1990, doi: 10.1523/jneurosci.10-02-00420.1990.

[32] T. Hafting, M. Fyhn, S. Molden, M. B. Moser, and E. I. Moser, “Microstructure of a spatial map in the entorhinal cortex,” *Nature*, vol. 436, no. 7052, 2005, doi: 10.1038/nature03721.

[33] S. Sreenivasan and I. Fiete, “Grid cells generate an analog error-correcting code for singularly precise neural computation,” *Nat. Neurosci.*, vol. 14, no. 10, 2011, doi: 10.1038/nn.2901.

[34] F. Sargolini *et al.*, “Conjunctive representation of position, direction, and velocity in entorhinal cortex,” *Science (80-.).*, vol. 312, no. 5774, 2006, doi: 10.1126/science.1125572.

[35] M. Franzius, H. Sprekeler, and L. Wiskott, “Slowness and sparseness lead to place, head-direction, and spatial-view cells,” *PLoS Comput. Biol.*, vol. 3, no. 8, 2007, doi: 10.1371/journal.pcbi.0030166.

[36] A. Finkelstein, D. Derdikman, A. Rubin, J. N. Foerster, L. Las, and N. Ulanovsky, “Three-dimensional head-direction coding in the bat brain,” *Nature*, vol. 517, no. 7533, 2014, doi: 10.1038/nature14031.

[37] W. de Cothi and H. J. Spiers, “Spatial Cognition: Goal-Vector Cells in the Bat

Hippocampus," *Current Biology*, vol. 27, no. 6. 2017, doi: 10.1016/j.cub.2017.01.061.

[38] I. A. Bachelder and A. M. Waxman, "Mobile robot visual mapping and localization: A view-based neurocomputational architecture that emulates hippocampal place learning," *Neural Networks*, vol. 7, no. 6–7, 1994, doi: 10.1016/S0893-6080(05)80160-1.

[39] T. Strösslin, R. Chavarriaga, D. Sheynikhovich, and W. Gerstner, "Modelling path integrator recalibration using hippocampal place cells," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2005, vol. 3696 LNCS, doi: 10.1007/11550822_9.

[40] J. Steckel and H. Peremans, "BatSLAM: Simultaneous Localization and Mapping Using Biomimetic Sonar," *PLoS One*, vol. 8, no. 1, 2013, doi: 10.1371/journal.pone.0054076.

[41] D. Vanderelst, J. Steckel, A. Boen, H. Peremans, and M. W. Holderied, "Place recognition using batlike sonar," *Elife*, vol. 5, 2016, doi: 10.7554/eLife.14188.

[42] L. Jakobsen and A. Surlykke, "Vespertilionid bats control the width of their biosonar sound beam dynamically during prey pursuit," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 31, 2010, doi: 10.1073/pnas.1006630107.

[43] S. Koul and T. K. Horiuchi, "Waypoint Path Planning with Synaptic-Dependent Spike Latency," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 66, no. 4, 2019, doi: 10.1109/TCSI.2018.2882818.

[44] R. Ollington and P. Vamplew, "Concurrent q-learning for autonomous

mapping and navigation," *Int. Conf. Comput. Intell. Robot. Auton. Syst.*, 2003.

[45] J. C. Tapson *et al.*, "Synthesis of neural networks for spatio-temporal spike pattern recognition and processing," *Front. Neurosci.*, no. 7 AUG, 2013, doi: 10.3389/fnins.2013.00153.

[46] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," 2015.

[47] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," 2016.

[48] E. Bisong and E. Bisong, "Google Colaboratory," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, 2019.

[49] G. Bin Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: A survey," *Int. J. Mach. Learn. Cybern.*, vol. 2, no. 2, 2011, doi: 10.1007/s13042-011-0019-y.

[50] J. Tapson and A. van Schaik, "Learning the pseudoinverse solution to network weights," *Neural Networks*, vol. 45, 2013, doi: 10.1016/j.neunet.2013.02.008.

[51] J. Mogdans and H. U. Schnitzler, "Range resolution and the possible use of spectral information in the echolocating bat, Eptesicus fuscus," *J.Acoust.Soc.Am.*, vol. 88, no. 2, 1990.